COMMUN. MATH. SCI.
Vol. 18, No. 3, pp. 635–661

FORWARD BACKWARD DOUBLY STOCHASTIC DIFFERENTIAL EQUATIONS AND THE OPTIMAL FILTERING OF DIFFUSION PROCESSES*

FENG BAO†, YANZHAO CAO‡, AND XIAOYING HAN§

Abstract. The connection between forward backward doubly stochastic differential equations and the optimal filtering problem is established without using the Zakai equation. The solutions of forward backward doubly stochastic differential equations are expressed in terms of a conditional law of a partially observed Markov diffusion process. It then follows that the adjoint time-inverse forward backward doubly stochastic differential equations govern the evolution of the unnormalized filtering density in the optimal filtering problem.

Key words. Forward backward doubly stochastic differential equations; optimal filtering problem; Feynman-Kac formula; Itô's formula; adjoint stochastic processes.

AMS subject classifications. 60H10; 60H30.

1. Introduction

The purpose of this paper is to study nonlinear filtering problems with systems of forward backward doubly stochastic differential equations. The aim of a nonlinear filtering problem is to determine the optimal estimate of the state of a noise-perturbed dynamical system given noisy observations on the dynamics. Some of the pioneering work on optimal filtering problems is due to Kallianpur and Striebel [22] and Zakai [34]. In particular, the Kallianpur-Striebel formula, which characterizes the conditional probability density function (PDF) of the state as the solution of a nonlinear stochastic partial differential equation (SPDE), provides a continuous time framework of the optimal filtering, while the approach proposed by Zakai leads to a linear stochastic integro-differential parabolic equation, referred to as Zakai's equation. Under strong regularity conditions it can be shown that the solution of Zakai's equation represents an unnormalized conditional density of the state process, which is also called the "filtering density". Fundamental theoretical research on optimal filtering problems can also be found in Kalman and Bucy [10,23], Kushner and Pardoux [25,29], Shiryaev [32] and Stratonovich [33], and studies on discrete nonlinear filter solvers can be found in [1–3, 7–9, 11, 14, 15, 17, 18, 20, 24, 26, 27]). The advantage of solving optimal filtering problems by SPDEs, such as Zakai's equation, lies in that it provides an "exact" solution for the filtering density . However, such methods have not been widely used by the science and engineering community due to their high complexity [5,12,13,19].

An alternative for deriving the unnormalized conditional density function is through the solution of a system of stochastic (ordinary) differential equations (SDEs) which consists of two SDEs, one standard SDE and one backward doubly stochastic differential equation

*Received: May 11, 2017; Accepted (in revised form): November 18, 2019. Communicated by Arnulf Jentzen.
The first author acknowledges the support by the Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research through FASTMath Institute and CompFUSE project. The first author is also partially supported by National Science Foundation under grant number DMS1720222. The second author is partially supported by National Science Foundation under grant number DMS1620150.
†Department of Mathematics, Florida State University, Tallahassee, Florida 32306, USA (bao@math.fsu.edu).
‡Department of Mathematics and Statistics, Auburn University, Auburn, Alabama 36849, USA (yzc0009@auburn.edu).
§Department of Mathematics and Statistics, Auburn University, Auburn, Alabama 36849, USA (xzh0003@auburn.edu).

(BDSDE). We refer to such a system as forward backward doubly stochastic differential equation (FBDSDE) system.

FBDSDE systems were first studied by Pardoux and Peng in [31], where the connection between FBDSDEs and certain parabolic-type *backward* SPDEs was established. Our recent work [4, 8] indicates that solving optimal filtering problems by FBDSDE systems may be computationally more efficient than doing that by SPDEs. The theoretical foundation of the FBDSDE approach is the fact that the solution of the BDSDE in an FBDSDE system is equivalent to the solution of its corresponding backward SPDE [8, 31]. However, since Zakai's equation is a forward SPDE, we need to invert the time index in the FBDSDE system to change the propagation direction of the BDSDE from backward to forward so that it is consistent with Zakai's equation. In this way, we obtain a time-inverse FBDSDE system that is equivalent to Zakai's equation.

The primary goal of this work is to establish a direct link between the optimal filtering problem and the FBDSDE system without using SPDEs. The procedure consists of two steps. In the first step, we establish the FBDSDE version of the Feynman-Kac formula for the optimal filtering problem. In this FBDSDE system, the forward SDE is simply the SDE for the state of the optimal filtering problem, and the BDSDE contains two noise terms (thus the name "doubly"): one counting for the backward nature of the equation while the other counting for the observation process which is a Brownian motion under an appropriate Girsanov transform. It is worthy noting that without the observation noise, the BDSDE is reduced to a BSDE (backward stochastic differential equation), whose solution is simply the BSDE version of the Feynman-Kac formula. On the other hand, the coefficient of observation noise resembles the coefficient of the multiplicative noise in the Zakai equation [34].

In the second step, we derive the adjoint BDSDE for the BDSDE corresponding to the Feynman-Kac formula and show that its solution solves the nonlinear filtering problem. This is done using the fact that the inner production between the BDSDE of the Feynman-Kac formula and its adjoint is a constant. Such a connection is similar to the relationship between the PDE of Feynman-Kac formula for a SDE and its adjoint, which is the Fokker Planck equation.

To the best of our knowledge, similar results have not been obtained before. The main difficulty of the direct derivation of the BDSDE filter is the lack of knowledge of the BDSDE form of the Feynman-Kac formula for the nonlinear filtering problem and the corresponding adjoint BDSDE. In this sense, our work contributes to the understanding of BDSDE theory in its own right.

The rest of this paper is organized as follows. In Section 2 we present mathematical formulations of the optimal filtering problem and provide a brief introduction to FBDSDEs. In Section 3 we establish the connection between FBDSDEs and the unnormalized conditional density function. Some closing remarks are given in Section 4.

2. Preliminaries

In this section, we introduce the mathematical formulation of the main topics of this paper – the optimal filtering problem and FBDSDEs.

2.1. The optimal filtering problem.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{W_t\}_{t>0}$ and $\{B_t\}_{t>0}$ be two mutually independent standard Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{P})$, with values in $\mathbb{R}^d$ and $\mathbb{R}^l$, respectively. Denote by $\mathcal{N}$ the class of $\mathbb{P}$-null sets of $\mathcal{F}$. For each $t \in [0, T]$, where $T > 0$, and any process η_t, let

$$\mathcal{F}^{\eta}_{s,t} := \sigma\{\eta_r - \eta_s : s \le r \le t\} \vee \mathcal{N}$$

be the σ-field generated by $\{\eta_r - \eta_s\}_{s \le r \le t}$. When $s = 0$, we write $\mathcal{F}^{\eta}_t = \mathcal{F}^{\eta}_{0,t}$ in short.

For an optimal filtering problem, we are given the stochastic differential system on $(\Omega, \mathcal{F}, \mathbb{P})$

$$\begin{cases} \mathrm{d}U_t = \mu_t(U_t)\mathrm{d}t + \rho_t \mathrm{d}W_t + \tilde{\rho}_t \mathrm{d}B_t, \\ \mathrm{d}V_t = h(U_t)\mathrm{d}t + \mathrm{d}B_t, \end{cases} \tag{2.1}$$

where $\{U_t \in \mathbb{R}^d : t \geq 0\}$ is the "state process" that describes the state of a dynamical system, and $\{V_t \in \mathbb{R}^l : t \geq 0\}$ is the "measurement process" which is the noise perturbed observations of the state U_t. Given an initial state U_0 with probability distribution $p_0(u)$ independent of W_t and B_t, the goal of the optimal filtering problem is to obtain the best estimate of $\phi(U_t)$ as the conditional expectation with respect to the measurement $\{V_r\}_{0 \leq r \leq t}$, where ϕ is a given test function.

Denote by $\mathcal{F}_t^V := \sigma\{V_r : 0 \leq r \leq t\}$ the σ-field generated by the measurement process from time 0 to t and denote by $\mathcal{M}_t$ the space of all $\mathcal{F}_t^V$-measurable and square integrable random variables at time t. The optimal filtering problem can be formulated mathematically as to find the conditional expectation

$$\mathbb{E}\left[\phi(U_t) \middle| \mathcal{F}_t^V\right] = \inf\left\{\mathbb{E}[|\phi(U_t) - \psi_t|^2] : \psi_t \in \mathcal{M}_t\right\}.$$

According to [21,22], the *optimal filter* is given by the well known Kallianpur–Striebel formula

$$\mathbb{E}\left[\phi(U_t) \middle| \mathcal{F}_t^V\right] = \frac{\int_{\mathbb{R}^d} \phi(u) p_t \mathrm{d}u}{\int_{\mathbb{R}^d} p_t \mathrm{d}u}, \tag{2.2}$$

where p_t is the *unnormalized filtering density*. In [34], Zakai showed that p_t satisfies the following SPDE (Zakai equation).

$$dp(t,x) = L^* p(t,x)dt + g(x)p(t,x)dV_t, \quad t > 0, x \in R^d, \tag{2.3}$$

where

$$L = \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^\top)_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i}.$$

We can see from (2.2) that finding the unnormalized filtering density p_t is equivalent to obtaining the optimal filter $\mathbb{E}\left[\phi(U_t) \middle| \mathcal{F}_t^V\right]$. In our BDSDE filter, we aim to derive a FBDSDE system to solve for p_t.

Define

$$Q_t^s := \exp\left\{\int_s^t h(U_r)\mathrm{d}U_r - \frac{1}{2}\int_s^t |h(U_r)|^2 \mathrm{d}r\right\}.$$

When $s = 0$, we denote Q_t^0 as Q_t in short. Let $\tilde{\mathbb{P}}$ be the probability measure induced on the space $(\Omega, \mathcal{F})$ such that

$$\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t^V} = Q_t. \tag{2.4}$$

Then according to the Cameron-Martin theorem the probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent when the Novikov condition is satisfied and the measurement process V is a standard Brownian motion under the induced probability $\tilde{\mathbb{P}}$ [16]. Moreover,

$$\mathbb{E}\left[\phi(U_t)\middle|\mathcal{F}_t^V\right]=\frac{\tilde{\mathbb{E}}\left[\phi(U_t)Q_t\middle|\mathcal{F}_t^V\right]}{\tilde{\mathbb{E}}\left[Q_t\middle|\mathcal{F}_t^V\right]}, \tag{2.5}$$

where $\tilde{\mathbb{E}}$ denotes the expectation with respect to $\tilde{\mathbb{P}}$ (see [28], Lemma 8.6.2).

2.2. FBDSDE systems. For each $t\in[0,T]$, define

$$\mathcal{F}_t:=\mathcal{F}_t^W\vee\mathcal{F}_{t,T}^B.$$

Note that the collection $\{\mathcal{F}_t:t\in[0,T]\}$ is neither increasing nor decreasing, and thus does not constitute a filtration [31]. For any positive integer $n\in\mathbb{N}$, denote by $\mathcal{M}^2(0,T;\mathbb{R}^n)$ the set of $\mathbb{R}^n$-valued jointly measurable random processes $\{\psi_t:t\in[0,T]\}$ such that ψ_t is $\mathcal{F}_t$ measurable for a.e. $t\in[0,T]$ and satisfies

$$\mathbb{E}\int_0^T|\psi_t|^2\mathrm{d}t<\infty.$$

Similarly, denote by $\mathcal{S}^2([0,T];\mathbb{R}^n)$ the set of continuous $\mathbb{R}^n$-valued random processes $\{\psi_t:t\in[0,T]\}$ such that ψ_t is $\mathcal{F}_t$ measurable for any $t\in[0,T]$ and satisfies

$$\mathbb{E}\sup_{0\le t\le T}|\psi_t|^2<\infty.$$

Next we provide a brief introduction to FBDSDE systems (see [31] for details). Given $\tau\ge 0$, $x\in\mathbb{R}^d$ and $\varphi\in\mathcal{L}^2(\Omega,\mathcal{F}_T,\mathbb{P})$, a FBDSDE system can be formulated as

$$\begin{cases}\mathrm{d}X_t=\mu(X_t)\mathrm{d}t+\sigma(X_t)\mathrm{d}W_t, & \tau\le t\le T,\\ -\mathrm{d}Y_t=f(t,X_t,Y_t,Z_t)\mathrm{d}t+g(t,X_t,Y_t,Z_t)\mathrm{d}\overleftarrow{B}_t-Z_t\mathrm{d}W_t, & \tau\le t\le T,\\ X_\tau=x, \quad Y_T=\varphi(X_T),\end{cases} \tag{2.6}$$

or, in the integral equation form, for any $t\in[\tau,T]$,

$$X_t=x+\int_\tau^t\mu(X_s)\mathrm{d}s+\int_\tau^t\sigma(X_s)\mathrm{d}W_s, \tag{2.7}$$

$$Y_t=\varphi(X_T)+\int_t^T f(s,X_s,Y_s,Z_s)\mathrm{d}s+\int_t^T g(s,X_s,Y_s,Z_s)\mathrm{d}\overleftarrow{B}_s-\int_t^T Z_s\mathrm{d}W_s. \tag{2.8}$$

Notice that Equation (2.7) is a standard forward SDE with a standard forward Itô integral and Equation (2.8) is a BDSDE involving the backward Itô integral $\int\cdot\mathrm{d}\overleftarrow{B}_s$ (see [30] for details on the two types of integrals).

Let the mappings $f:[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}\to\mathbb{R}^k$ and $g:[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}\to\mathbb{R}^{k\times l}$ be jointly measurable and for any $(y,z)\in\mathbb{R}^k\times\mathbb{R}^{k\times d}$,

$$f(\cdot,\cdot,y,z)\in\mathcal{M}^2(0,T;\mathbb{R}^k),\qquad g(\cdot,\cdot,y,z)\in\mathcal{M}^2(0,T;\mathbb{R}^{k\times l}).$$

Denote by $|\cdot|$ the Euclidean norm of a vector and by $\|A\|:=\sqrt{\mathrm{Tr}(AA^*)}$ the norm of a matrix A. The existence and uniqueness of solutions, moment estimates for the solutions, and the regularity of solutions to Equation (2.8) rely on the following assumptions.

ASSUMPTION 2.1. *f and g satisfy the Lipschitz condition: there exist constants $c>0$ and $0<\bar{c}<1$ such that for any $(t,x)\in[0,T]\times\mathbb{R}^d$, $y_1,y_2\in\mathbb{R}^k$ and $z_1,z_2\in\mathbb{R}^{k\times d}$,*

$$|f(t,x,y_1,z_1)-f(t,x,y_2,z_2)|^2\le c(|y_1-y_2|^2+\|z_1-z_2\|^2),$$
$$\|g(t,x,y_1,z_1)-g(t,x,y_2,z_2)\|^2\le c|y_1-y_2|^2+\bar{c}\|z_1-z_2\|^2.$$

ASSUMPTION 2.2. *There exists $c>0$ such that for all $(t,x,y,z)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$,*

$$gg^*(t,x,y,z)\le zz^*+c(\|g(t,x,0,0)\|^2+|y|^2)I.$$

ASSUMPTION 2.3. *For any $(t,x,y,z)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}$ and $\theta\in\mathbb{R}^{k\times d}$*

$$\frac{\partial g}{\partial z}(t,x,y,z)\theta\theta^*\left(\frac{\partial g}{\partial z}(t,x,y,z)\right)^*\le\theta\theta^*.$$

The following results are due to Pardoux and Peng [31].

PROPOSITION 2.1. *Under Assumption 2.1, the BDSDE (2.8) admits a unique solution*

$$(Y,Z)\in\mathcal{S}^2([0,T];\mathbb{R}^k)\times\mathcal{M}^2(0,T;\mathbb{R}^{k\times d}).$$

For any positive integer k, denote by $\mathcal{C}_b^k(\mathbb{R}^m;\mathbb{R}^n)$ the collection of $\mathcal{C}^k$ functions from $\mathbb{R}^m$ into $\mathbb{R}^n$ with bounded partial derivatives of all orders less than or equal to k, and denote by $\mathcal{C}_p^k(\mathbb{R}^m;\mathbb{R}^n)$ the collection of $\mathcal{C}^k$ functions from $\mathbb{R}^m$ into $\mathbb{R}^n$ with partial derivatives of all orders less than or equal to k which grow at most like a polynomial function of x as $x\to\infty$. It is well known that given $\mu\in\mathcal{C}_b^3(\mathbb{R}^d;\mathbb{R}^d)$ and $\sigma\in\mathcal{C}_b^3(\mathbb{R}^d;\mathbb{R}^{d\times d})$, for each $(\tau,x)\in[0,T]\times\mathbb{R}^d$, the SDE (2.7) has a unique strong solution, denoted as $X_t^{\tau,x}$. Consequently we also denote by $(Y_t^{\tau,x},Z_t^{\tau,x})$ the unique solution to the BDSDE

$$Y_t=\varphi(X_T^{\tau,x})+\int_t^T f(s,X_s^{\tau,x},Y_s,Z_s)\mathrm{d}s+\int_t^T g(s,X_s^{\tau,x},Y_s,Z_s)\mathrm{d}\overleftarrow{B}_s-\int_t^T Z_s\mathrm{d}W_s. \tag{2.9}$$

PROPOSITION 2.2. *Let $\varphi\in\mathcal{C}_p^3(\mathbb{R}^d;\mathbb{R}^k)$. Under Assumptions 2.1–2.3, the random field $\{Y_\tau^{\tau,x}:\tau\in[0,T],x\in\mathbb{R}^d\}$ admits a continuous version such that for any $\tau\in[0,T]$, $x\mapsto Y_\tau^{\tau,x}$ is of class $\mathcal{C}^2$ a.s..*

In addition to the existence and uniqueness result in Proposition 2.1, it is also worthy to notice that the Assumptions 2.1–2.3 allow a proof of the existence of L^p solutions for $p>2$.

The following regularity result can be obtained with techniques (see Proposition 1 in [6]).

LEMMA 2.1. *In addition to Assumption 2.1, assume that $f,g\in\mathcal{C}_b^1$. Then the solution $(Y_t^{\tau,x},Z_t^{\tau,x})$ to the BDSDE (2.9) satisfies*

$$\mathbb{E}\left[(Y_t^{\tau,x}-Y_\tau^{\tau,x})^2\right]\le C(t-\tau),\quad \mathbb{E}\left[(Z_t^{\tau,x}-Z_\tau^{\tau,x})^2\right]\le C(t-\tau),\quad 0\le\tau\le t\le T,$$

where C is a positive constant independent of τ and t.

Note that with the convention above, the unique solution to the FBDSDE system (2.7) – (2.8) can be written as $(X_t^{\tau,x},Y_t^{\tau,x},Z_t^{\tau,x})$. Denote

$$\nabla X_t^{\tau,x}:=\frac{\partial X_t^{\tau,x}}{\partial x},\qquad \nabla Y_t^{\tau,x}:=\frac{\partial Y_t^{\tau,x}}{\partial x},\qquad \nabla Z_t^{\tau,x}:=\frac{\partial Z_t^{\tau,x}}{\partial x}.$$

Then $(\nabla Y_t^{\tau,x}, \nabla Z_t^{\tau,x})$ is the unique solution to variational form of the BDSDE (2.8) (see [31])

$$\nabla Y_t^{\tau,x} = \varphi'(X_T^{\tau,x})\nabla X_T^{\tau,x} + \int_t^T \left(\frac{\partial f}{\partial x}\nabla X_s^{\tau,x} + \frac{\partial f}{\partial Y}\nabla Y_s^{\tau,x} + \frac{\partial f}{\partial Z}\nabla Z_s^{\tau,x}\right)\mathrm{d}s$$
$$+ \int_t^T \left(\frac{\partial g}{\partial x}\nabla X_s^{\tau,x} + \frac{\partial g}{\partial Y}\nabla Y_s^{\tau,x} + \frac{\partial g}{\partial Z}\nabla Z_s^{\tau,x}\right)\mathrm{d}\overleftarrow{B}_s - \int_t^T \nabla Z_s^{\tau,x}\mathrm{d}W_s.$$

In addition, the random field $\left\{Z_t^{\tau,x} : t \in [\tau,T], x \in \mathbb{R}^d\right\}$ has an a.s. continuous version

$$Z_t^{\tau,x} = \nabla Y_t^{\tau,x}(\nabla X_t^{\tau,x})^{-1}\sigma(X_t^{\tau,x}), \qquad Z_\tau^{\tau,x} = \nabla Y_\tau^{\tau,x}\sigma(x). \tag{2.10}$$

The following Lemma follows directly from Proposition 2.1 and Lemma 2.1.

LEMMA 2.2. *Assume that $\mu \in \mathcal{C}_b^2$, $f \in \mathcal{C}_b^2$, $g \in \mathcal{C}_b^2$ and $\varphi \in \mathcal{C}_b^2$. Then there exists $C > 0$ such that*

$$\mathbb{E}[(\nabla Y_t^{\tau,x} - \nabla Y_t^{\tau,x})^2] \le C(t-\tau), \quad \mathbb{E}[(\nabla Z_t^{\tau,x} - \nabla Z_\tau^{\tau,x})^2] \le C(t-\tau), \quad 0 \le \tau \le t \le T.$$

Moreover,

$$\mathbb{E}\sup_{0\le t\le T} |\nabla Y_t^{\tau,x}|^2 < \infty.$$

3. FBDSDEs and optimal filtering

In this section, we show that the solution of an optimal filtering problem can be obtained by solving a FBDSDE system. To this end, we first prove a Feynman-Kac type formula in the form of the first FBDSDE system defined in (2.6) without the deterministic integral $f(t, X_t, Y_t, Z_t)\mathrm{d}t$. Then we derive a FBDSDE system which is the adjoint of the FBDSDE system in the Feynman Kac formula and prove that its solution provides an unnormalized filtering density sought in the optimal filtering problem defined in Section 3.3. For simplicity of exposition, we only discuss the one dimensional case with $d = 1$ and $l = 1$. The methodology developed can be easily applied to multi-dimensional cases.

3.1. Feynman-Kac type formula for optimal filtering. For $\tau \in [0,T]$ and $x \in \mathbb{R}^d$, consider the following FBDSDE system on the probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$

$$\begin{cases} \mathrm{d}X_t = \mu_t(X_t)\mathrm{d}t + \sigma_t \mathrm{d}W_t, \quad \tau \le t \le T & \text{(SDE)} \\ -\mathrm{d}Y_t = -Z_t\mathrm{d}W_t + \left(h(X_t)Y_t + \dfrac{\tilde{\rho}_t}{\sigma_t}Z_t\right)\mathrm{d}\overleftarrow{V}_t, \quad \tau \le t \le T & \text{(BDSDE)} \\ X_\tau = x, \qquad Y_T = \phi(X_T), \end{cases} \tag{3.1}$$

where $\sigma_t^2 = \rho_t^2 + \tilde{\rho}_t^2$, and μ, ρ, $\tilde{\rho}$, h are the functions appearing in the optimal filtering problem (2.1). Here W_t is the same Brownian motion as in the nonlinear filtering problem (2.1), while V_t is the measurement process which becomes a standard Brownian motion independent of W_t under the induced probability measure $\tilde{\mathbb{P}}$ defined by (2.4). Then X_t is a $\mathcal{F}^W$ adapted stochastic process and the pair (Y_t, Z_t) is adapted to $\mathcal{F}_t^W \vee \mathcal{F}_{t,T}^V$. For any single-variable function $f = f(x)$, we denote $f' := \frac{\mathrm{d}f}{\mathrm{d}x}$ and $f'' = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$.

Next, we show that the FBDSDE system (3.1) provides the Feynman-Kac formula in the optimal filtering context. But first we state a couple of remarks.

REMARK 3.1. Without the observation V_t, the BDSDE in (3.1) is reduced to a simple BSDE, whose exact solution is given by

$$Y(t) = E[\phi(X_T)|\mathcal{F}^W(t), X(\tau) = x], \quad t \ge \tau.$$

Thus

$$Y(\tau)=E[\phi(X_T)|X(\tau)=x]$$

which is the standard Feynman-Kac formula.

REMARK 3.2. On the other hand, when the observation V_t does exist, the coefficient of dV_t in (3.1) resembles the coefficient of multiplicative noise term in the Zakai equation [34] with $\frac{Z_t}{\sigma_t}$ in the BSDEs being replaced by ∇u_t in the Zakai equation.

THEOREM 3.1. *Assume that ϕ is bounded, $\mu_t,\rho_t,\tilde{\rho}_t\in\mathcal{C}_b(\mathbb{R};\mathbb{R})$ and $h\in\mathcal{C}_b^2(\mathbb{R};\mathbb{R})$. Then, $\forall\tau\in[0,T]$ and $x\in\mathbb{R}^d$ the following equality holds a.s.*

$$Y_\tau^{\tau,x}=\mathbb{E}_\tau^x[\phi(U_T)Q_T^\tau],\quad \textit{with}\quad \mathbb{E}_\tau^x[\cdot]:=\tilde{\mathbb{E}}[\cdot|\mathcal{F}_{\tau,T}^V,U_\tau=x]. \tag{3.2}$$

In order to prove Theorem 3.1 , we first introduce a regularity lemma as follows.

LEMMA 3.1. *Assume that μ_t and σ_t are bounded and $h\in\mathcal{C}_b^2(\mathbb{R};\mathbb{R})$. Then for any $0\le s\le t\le T$, there exists a positive constant C independent of s and t such that*

$$\tilde{\mathbb{E}}[(h(X_t)-h(X_s))^2|\mathcal{F}_{t,T}^V]\le C(t-s). \tag{3.3}$$

Proof. The application of Itô's formula to $h(X_t)$ results in

$$h(X_t)=h(X_s)+\int_s^t\left(\mu_r(X_r)h'(X_r)+\frac{\sigma_r^2}{2}h''(X_r)\right)\mathrm{d}r+\int_s^t\sigma_r h'(X_r)\mathrm{d}W_r,$$

and hence

$$(h(X_t)-h(X_s))^2=\left(\int_s^t\left(\mu_r(X_r)h'(X_r)+\frac{\sigma_r^2}{2}h''(X_r)\right)\mathrm{d}r+\int_s^t\sigma_r h'(X_r)\mathrm{d}W_r\right)^2. \tag{3.4}$$

Taking expectation $\tilde{\mathbb{E}}$ of (3.4) gives

$$\tilde{\mathbb{E}}\big[(h(X_t)-h(X_s))^2\big]=\tilde{\mathbb{E}}\left[\left(\int_s^t(\mu_r(X_r)h'(X_r)+\frac{\sigma_r^2}{2}h''(X_r))\mathrm{d}r\right)^2\right]+\tilde{\mathbb{E}}\left[\int_s^t(\sigma_r h'(X_r))^2\mathrm{d}r\right].$$

The inequality (3.3) then follows immediately from the assumptions of the lemma. ∎

With Proposition 2.1, Lemma 2.1 and Lemma 3.1, we are ready to prove Theorem 3.1.

Proof. **(Proof of Theorem 3.1.)** We prove the statement (3.2) for $\tau=0$ only, the general case follows from the $\tau=0$ case trivially. It is straightforward to verify that under assumptions in Theorem 3.1, all the assumptions of Proposition 2.1, and Lemmas 2.1 and 3.1 are fulfilled. Since $Y_\tau^{\tau,x}$ and $Z_\tau^{\tau,x}$ are functions of x, we write $Y_\tau^{\tau,x}=Y_\tau(x)$ and $Z_\tau^{\tau,x}=Z_\tau(x)$ in the sequel.

Let $0=t_0<t_1<t_2\cdots<t_N=T$ be a uniform partition of $[0,T]$ with $t_{n+1}-t_n=T/N:=\Delta t$ and define

$$\Delta_n=\mathbb{E}_0^x[Q_{t_{n+1}}Y_{t_{n+1}}(U_{t_{n+1}})-Q_{t_n}Y_{t_n}(U_{t_n})].$$

It follows immediately that

$$\mathbb{E}_0^x[\phi(U_T)Q_T-Y_0(x)]=\sum_{n=0}^{N-1}\Delta_n.$$

Denote $\tilde{\mathbb{P}}_x := \tilde{\mathbb{P}}(\cdot|U_0 = x)$. To prove (3.2) it suffices to verify that

$$\sum_{n=0}^{N-1} \Delta_n \overset{N\to\infty}{\longrightarrow} 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.5}$$

For each $n \geq 0$, let U_{t_n} be the state in (2.1) at time step t_n and consider the FBDSDE system (3.1) on $[t_n, t_{n+1}]$ with initial condition U_{t_n}:

$$\begin{cases} \mathrm{d}\hat{X}_t = \mu_t(\hat{X}_t)\mathrm{d}t + \sigma_t \mathrm{d}W_t, \\ -\mathrm{d}Y_t = -Z_t \mathrm{d}W_t + \left(h(\hat{X}_t)Y_t + \dfrac{\tilde{\rho}_t}{\sigma_t} Z_t \right) \mathrm{d}\overleftarrow{V}_t, \\ \hat{X}_{t_n} = U_{t_n}, \quad Y_{t_{n+1}} = Y_{t_{n+1}}(\hat{X}_{t_{n+1}}), \end{cases} \tag{3.6}$$

where the notation $\hat{X}_t$ is introduced to emphasize the initial condition $\hat{X}_{t_n} = U_{t_n}$, which incorporates the state process U_t into the FBDSDE system and gives us the identity $Y_{t_n}(U_{t_n}) = Y_{t_n}(\hat{X}_{t_n})$. However, despite the equality of U_{t_n} and $\hat{X}_{t_n}$, from the definition of U_t in (2.1) and $\hat{X}_t$ in (3.6), we still observe that there's a difference between $U_{t_{n+1}}$ and $\hat{X}_{t_{n+1}}$, which is described as the following

$$U_{t_{n+1}} = \hat{X}_{t_{n+1}} + \int_{t_n}^{t_{n+1}} \rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}} \sigma_s \mathrm{d}W_s + \int_{t_n}^{t_{n+1}} \tilde{\rho}_s (\mathrm{d}V_s - h(U_s)\mathrm{d}s) + R_X^{n+1},$$

where $R_X^{n+1} = \int_{t_n}^{t_{n+1}} \mu_s(U_s)\mathrm{d}s - \int_{t_n}^{t_{n+1}} \mu_s(\hat{X}_s)\mathrm{d}s$. To simplify presentation, for any process ψ_t we write $\hat{\psi}_t := \psi_t(\hat{X}_t)$ throughout the rest of this proof and let $\eta_{n+1} := U_{t_{n+1}} - \hat{X}_{t_{n+1}}$ be the difference of $U_{t_{n+1}}$ and $\hat{X}_{t_{n+1}}$. Then the above equation becomes

$$\eta_{n+1} = \int_{t_n}^{t_{n+1}} \rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}} \sigma_s \mathrm{d}W_s + \int_{t_n}^{t_{n+1}} \tilde{\rho}_s (\mathrm{d}V_s - h(U_s)\mathrm{d}s) + R_X^{n+1}. \tag{3.7}$$

Applying the Taylor expansion to $Y_{t_{n+1}}$ we have that

$$Y_{t_{n+1}}(U_{t_{n+1}}) = \hat{Y}_{t_{n+1}} + \hat{Y}'_{t_{n+1}} \cdot \eta_{n+1} + \frac{1}{2}\hat{Y}''_{t_{n+1}} \cdot (\eta_{n+1})^2 + \xi_{n+1}, \tag{3.8}$$

where ξ_{n+1} is the Taylor remainder such that $\mathbb{E}_0^x[(\xi_{n+1})^2] \leq C(\Delta t)^3$. To deal with the expansion terms in (3.8) caused by the difference between $U_{t_{n+1}}$ and $\hat{X}_{t_{n+1}}$, we rewrite the expression for Δ_n as

$$\begin{aligned} \Delta_n &= \mathbb{E}_0^x \left[Q_{t_{n+1}} Y_{t_{n+1}}(U_{t_{n+1}}) - Q_{t_n} \hat{Y}_{t_{n+1}} + Q_{t_n} \hat{Y}_{t_{n+1}} - Q_{t_n} Y_{t_n}(U_{t_n}) \right] \\ &= \underbrace{\mathbb{E}_0^x \left[\left(Q_{t_{n+1}} - Q_{t_n} \right) \hat{Y}_{t_{n+1}} \right]}_{\textbf{(i)}} + \underbrace{\mathbb{E}_0^x \left[Q_{t_n} \left(\hat{Y}_{t_{n+1}} - Y_{t_n}(U_{t_n}) \right) \right]}_{\textbf{(ii)}} \\ &\qquad + \underbrace{\mathbb{E}_0^x \left[Q_{t_{n+1}} \left(\hat{Y}'_{t_{n+1}} \eta_{n+1} + \frac{1}{2}\hat{Y}''_{t_{n+1}} \cdot (\eta_{n+1})^2 + \xi_{n+1} \right) \right]}_{\textbf{(iii)}}. \end{aligned} \tag{3.9}$$

The rest of this proof can be divided into two parts. The first part is to estimate terms (**i**), (**ii**) and (**iii**) in (3.9) one by one and try to split Δ_n into several terms without $\hat{Y}$ derivatives.

Then, in the second part, we prove that the derived terms in Δ_n from the first part will converge to 0 in $\mathcal{L}^1$ after the summation from 0 to $N-1$.

Part I.

(**i**) Write $h_t = h(U_t)$ and $\hat{h}_t = h(\hat{X}_t)$, and applying Ito's formula to Q_{t_n} we obtain

$$\begin{aligned}\mathbb{E}_0^x\left[(Q_{t_{n+1}}-Q_{t_n})\hat{Y}_{t_{n+1}}\right] &= \mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}} h_s Q_s \mathrm{d}V_s \hat{Y}_{t_{n+1}}\right] \\ &= \mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}} \hat{h}_s Q_s \mathrm{d}V_s \hat{Y}_{t_{n+1}}\right] + \mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}} (h_s-\hat{h}_s) Q_s \mathrm{d}V_s \hat{Y}_{t_{n+1}}\right].\end{aligned} \tag{3.10}$$

Applying Itô's formula to function h yields

$$h_s - \hat{h}_s = h'(U_{t_n})\left(\int_{t_n}^s \rho_r \mathrm{d}W_r + \int_{t_n}^s \tilde{\rho}_r \mathrm{d}V_r - \int_{t_n}^s \sigma_r \mathrm{d}W_r\right) + \mathcal{O}(\Delta t),$$

and consequently with $h'_{t_n} := h'(U_{t_n})$ we have

$$\begin{aligned}&\mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}} (h_s-\hat{h}_s) Q_s \mathrm{d}V_s \hat{Y}_{t_{n+1}}\right] \\ =&\mathbb{E}_0^x\left[h'_{t_n} Q_{t_n} Y_{t_n}(U_{t_n}) \int_{t_n}^{t_{n+1}} \mathrm{d}V_s \left(\int_{t_n}^s \rho_r \mathrm{d}W_r - \int_{t_n}^s \sigma_r \mathrm{d}W_r\right)\right] \\ &+ \mathbb{E}_0^x\left[h'_{t_n}(Q_s - Q_{t_n})\left(\hat{Y}_{t_{n+1}} - Y_{t_n}(U_{t_n})\right) \int_{t_n}^{t_{n+1}} \mathrm{d}V_s \left(\int_{t_n}^s \rho_r \mathrm{d}W_r - \int_{t_n}^s \sigma_r \mathrm{d}W_r\right)\right] \\ &+ \mathbb{E}_0^x\left[h'_{t_n} Q_{t_n} \hat{Y}_{t_{n+1}} \int_{t_n}^{t_{n+1}} \int_{t_n}^s \tilde{\rho}_r \mathrm{d}V_r \mathrm{d}V_s\right] + \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right).\end{aligned} \tag{3.11}$$

First noting that $h'_{t_n} Q_{t_n} Y_{t_n}(U_{t_n}) \int_{t_n}^{t_{n+1}} \mathrm{d}V_s$ is independent of $\int_{t_n}^s \rho_r \mathrm{d}W_r - \int_{t_n}^s \sigma_r \mathrm{d}W_r$, we have

$$\mathbb{E}_0^x\left[h'_{t_n} Q_{t_n} Y_{t_n}(U_{t_n}) \int_{t_n}^{t_{n+1}} \mathrm{d}V_s \left(\int_{t_n}^s \rho_r \mathrm{d}W_r - \int_{t_n}^s \sigma_r \mathrm{d}W_r\right)\right] = 0. \tag{3.12}$$

Second, it's straightforward to verify that

$$\mathbb{E}_0^x\left[h'_{t_n}(Q_s - Q_{t_n})\left(\hat{Y}_{t_{n+1}} - Y_{t_n}(U_{t_n})\right) \int_{t_n}^{t_{n+1}} \mathrm{d}V_s \left(\int_{t_n}^s \rho_r \mathrm{d}W_r - \int_{t_n}^s \sigma_r \mathrm{d}W_r\right)\right] \sim \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right). \tag{3.13}$$

Putting (3.12) and (3.13) in (3.11), it follows from the regularity condition of $\tilde{\rho}_r$ that

$$\mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}} (h_s-\hat{h}_s) Q_s \mathrm{d}V_s \hat{Y}_{t_{n+1}}\right] = \mathbb{E}_0^x\left[h'_{t_n} Q_{t_n} \hat{Y}_{t_{n+1}} \tilde{\rho}_{t_n} \int_{t_n}^{t_{n+1}} \int_{t_n}^s \mathrm{d}V_r \mathrm{d}V_s\right] + \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right). \tag{3.14}$$

Define

$$\nu_n := h'_{t_n} Q_{t_n} \hat{Y}_{t_{n+1}} \tilde{\rho}_{t_n} \int_{t_n}^{t_{n+1}} \int_{t_n}^s \mathrm{d}V_r \mathrm{d}V_s. \tag{3.15}$$

Then by using the facts $\int_{t_n}^{t_{n+1}} \int_{t_n}^s \mathrm{d}V_r \mathrm{d}V_s = \frac{1}{2}\left((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\right)$ and $h'_{t_n} Q_{t_n} \hat{Y}_{t_{n+1}} \tilde{\rho}_{t_n}$ is independent of $\frac{1}{2}\left((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\right)$ we have

$$\sum_{n=0}^{N-1} \nu_n = \sum_{n=0}^{N-1} h'_{t_n} Q_{t_n} \hat{Y}_{t_{n+1}} \tilde{\rho}_{t_n} \cdot \frac{1}{2}\left((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\right) \overset{N\to\infty}{\longrightarrow} 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.16}$$

In summary (3.10) and (3.14)–(3.16) together give the estimate of the term (**i**) in (3.9) as

$$(\mathbf{i}) = \mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}} \hat{h}_s Q_s \mathrm{d}V_s \hat{Y}_{t_{n+1}}\right] + \mathbb{E}_0^x[\nu_n] + \mathcal{O}\left((\Delta)^{\frac{3}{2}}\right), \tag{3.17}$$

with $\sum_{n=0}^{N-1}\mathbb{E}_0^x[\nu_n] \to 0$ in $\mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x)$ as $N\to\infty$.

(**ii**) It follows directly from the FBDSDE system (3.6) that term (**ii**) in (3.9) satisfies

$$\begin{aligned}(\mathbf{ii}) &= \mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}} \hat{Z}_s \mathrm{d}W_s - Q_{t_n}\int_{t_n}^{t_{n+1}}\left(\hat{h}_s\hat{Y}_s + \frac{\tilde{\rho}_s}{\sigma_s}\hat{Z}_s\right)\mathrm{d}\overleftarrow{V}_s\right] \\ &= -\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\left(\hat{h}_s\hat{Y}_s + \frac{\tilde{\rho}_s}{\sigma_s}\hat{Z}_s\right)\mathrm{d}\overleftarrow{V}_s\right].\end{aligned} \tag{3.18}$$

(**iii**) Now, we are going to estimate the Taylor expansion terms in (3.9). By splitting term (**iii**) in (3.9) and using the definition of η_{n+1} in (3.7) we obtain

$$\begin{aligned}(\mathbf{iii}) &= \mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}\eta_{n+1}\right] + \frac{1}{2}\mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}''_{t_{n+1}}\cdot(\eta_{n+1})^2\right] + \mathbb{E}_0^x\left[Q_{t_{n+1}}\xi_{n+1}\right] \\ &= \underbrace{\mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\right]}_{(\mathbf{iii-1})} + \underbrace{\mathbb{E}_0^x\left[Q_{t_n}\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s \mathrm{d}V_s\right]}_{(\mathbf{iii-2})} \\ &\quad + \underbrace{\mathbb{E}_0^x\left[(Q_{t_{n+1}} - Q_{t_n})\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s \mathrm{d}V_s\right] - \mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s h_s \mathrm{d}s\right]}_{(\mathbf{iii-3})} \\ &\quad + \underbrace{\mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}R_X^{n+1}\right]}_{(\mathbf{iii-4})} + \frac{1}{2}\underbrace{\mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}''_{t_{n+1}}\cdot(\eta_{n+1})^2\right]}_{(\mathbf{iii-5})} + \mathbb{E}_0^x\left[Q_{t_{n+1}}\xi_{n+1}\right].\end{aligned} \tag{3.19}$$

We next estimate terms (**iii-1**) – (**iii-5**).

Denote

$$\nabla\hat{X}_t := \left.\frac{\partial \hat{X}_t^{t_n,x}}{\partial x}\right|_{x=U_{t_n}}, \qquad \nabla\hat{Y}_t := \left.\frac{\partial Y_t^{t_n,x}}{\partial x}\right|_{x=U_{t_n}}, \qquad \nabla\hat{Z}_t := \left.\frac{\partial Z_t^{t_n,x}}{\partial x}\right|_{x=U_{t_n}}.$$

Then term (**iii-1**) can be written as

$$(\mathbf{iii-1}) = \mathbb{E}_0^x\left[Q_{t_{n+1}}\left(\hat{Y}'_{t_{n+1}}\nabla\hat{X}_{t_{n+1}}\right)\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\left(\nabla\hat{X}_{t_{n+1}}\right)^{-1}\right]. \tag{3.20}$$

By using the fact $|(\nabla\hat{X}_{t_{n+1}})^{-1}| = 1 + \mathcal{O}(\Delta t)$ and the variational equation (see [31])

$$\nabla\hat{Y}_t = \hat{Y}'_{t_{n+1}}\nabla\hat{X}_{t_{n+1}} + \int_t^{t_{n+1}}\left(\hat{h}'_s\hat{Y}_s\nabla\hat{X}_s + \hat{h}_s\nabla\hat{Y}_s + \frac{\tilde{\rho}_s}{\sigma_s}\nabla\hat{Z}_s\right)\mathrm{d}\overleftarrow{V}_s - \int_t^{t_{n+1}}\nabla\hat{Z}_s \mathrm{d}W_s,$$

we deduce that (3.20) becomes

$$
\begin{aligned}
(\mathbf{iii-1}) = &\mathbb{E}_0^x\left[Q_{t_{n+1}}\left(\hat{Y}'_{t_{n+1}}\nabla\hat{X}_{t_{n+1}} - \nabla\hat{Y}_{t_n}\right)\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\right] \\
&+ \mathbb{E}_0^x\left[Q_{t_{n+1}}\nabla\hat{Y}_{t_n}\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\right] + \mathcal{O}((\Delta t)^{\frac{3}{2}}) \\
= &\mathbb{E}_0^x\left[\left(Q_{t_n}\int_{t_n}^{t_{n+1}}\nabla\hat{Z}_s \mathrm{d}W_s + \lambda_{t_n}\right)\cdot\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\right] + \mathcal{O}((\Delta t)^{\frac{3}{2}}),
\end{aligned}
$$

where $\lambda_{t_n} = -Q_{t_n}\int_{t_n}^{t_{n+1}}\left(\hat{h}'_s\bar{Y}_s\nabla\hat{X}_s + \hat{h}_s\nabla\hat{Y}_s + \frac{\tilde{\rho}_s}{\sigma_s}\nabla\hat{Z}_s\right)d\overleftarrow{V}_s + Q_{t_{n+1}}\nabla\hat{Y}_{t_n}$ is independent of $\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s$ and hence gives

$$
\mathbb{E}_0^x\left[\lambda_{t_n}\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\right] = 0.
$$

As a consequence

$$
\begin{aligned}
(\mathbf{iii-1}) = &\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\nabla\hat{Z}_s \mathrm{d}W_s\cdot\left(\int_{t_n}^{t_{n+1}}\rho_s \mathrm{d}W_s - \int_{t_n}^{t_{n+1}}\sigma_s \mathrm{d}W_s\right)\right] + \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right) \\
= &\mathbb{E}_0^x\left[Q_{t_n}\nabla\hat{Z}_{t_{n+1}}\cdot\int_{t_n}^{t_{n+1}}(\rho_s - \sigma_s)\mathrm{d}s\right] + \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right).
\end{aligned}
\tag{3.21}
$$

Let C represent a generic constant while the context is clear. By the definition of R_X^{n+1}, it is straightforward to verify that

$$
(\mathbf{iii-4}) = \mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}R_X^{n+1}\right] \le C(\Delta t)^{\frac{3}{2}}. \tag{3.22}
$$

Applying Itô's formula to Q_{t_n} in term (**iii-3**) we obtain

$$
\begin{aligned}
(\mathbf{iii-3}) = &\mathbb{E}_0^x\left[(Q_{t_{n+1}} - Q_{t_n})\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s \mathrm{d}V_s\right] - \mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s h_s \mathrm{d}s\right] \\
= &\mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}}h_s Q_s \mathrm{d}V_s\int_{t_n}^{t_{n+1}}\tilde{\rho}_s \mathrm{d}V_s\hat{Y}'_{t_{n+1}}\right] - \mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s h_s \mathrm{d}s\right] \\
\le &\left|\mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}}h_s(Q_s - Q_{t_{n+1}})\mathrm{d}V_s\int_{t_n}^{t_{n+1}}\tilde{\rho}_s \mathrm{d}V_s\hat{Y}'_{t_{n+1}}\right]\right| \\
&+ \left|\mathbb{E}_0^x\left[Q_{t_{n+1}}\hat{Y}'_{t_{n+1}}\left(\int_{t_n}^{t_{n+1}}h_s \mathrm{d}V_s\int_{t_n}^{t_{n+1}}\tilde{\rho}_s \mathrm{d}V_s - \int_{t_n}^{t_{n+1}}\tilde{\rho}_s h_s \mathrm{d}s\right)\right]\right| \\
\le &C(\Delta t)^{\frac{3}{2}}.
\end{aligned}
\tag{3.23}
$$

By using the definition of η_{n+1} in (3.7), we deduce that

$$
\begin{aligned}
(\mathbf{iii-5}) = &\frac{1}{2}\mathbb{E}_0^x\left[Q_{t_n}\hat{Y}''_{t_{n+1}}\int_{t_n}^{t_{n+1}}\left(\rho_s^2 + \tilde{\rho}_s^2 + \sigma_s^2 - 2\rho_s\sigma_s\right)\mathrm{d}s\right] + \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right) \\
= &\mathbb{E}_0^x\left[Q_{t_n}\hat{Y}''_{t_{n+1}}\int_{t_n}^{t_{n+1}}\left(\sigma_s^2 - \rho_s\sigma_s\right)\mathrm{d}s\right] + \mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right).
\end{aligned}
$$

As a simple corollary of the assertion (2.10), we have $\hat{Y}''_{t_{n+1}}\sigma_{t_{n+1}}=\nabla\hat{Z}_{t_{n+1}}+\mathcal{O}(\Delta t)$ and thus

$$(\mathbf{iii-5})=\mathbb{E}_0^x\left[Q_{t_n}\nabla\hat{Z}_{t_{n+1}}\int_{t_n}^{t_{n+1}}(\sigma_s-\rho_s)\mathrm{d}s\right]+\mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right). \tag{3.24}$$

It then remains to estimate term (**iii-2**). Notice that due to equations (2.10) and (3.1) we have $\hat{Z}_s/\sigma_s=\nabla\hat{Y}_s(\nabla\hat{X}_s)^{-1}$. Hence for any $s\in[t_n,t_{n+1}]$ it holds

$$\begin{aligned}\hat{Y}'_{t_{n+1}}-\frac{\hat{Z}_s}{\sigma_s}&=\hat{Y}'_{t_{n+1}}-\nabla\hat{Y}_s(\nabla\hat{X}_s)^{-1}\\&=-\int_s^{t_{n+1}}\left(\hat{h}'_r\hat{Y}_r\nabla\hat{X}_r+\hat{h}_r\nabla\hat{Y}_r+\frac{\tilde{\rho}_r}{\sigma_r}\nabla\hat{Z}_r\right)\mathrm{d}\overleftarrow{V}_r-\int_s^{t_{n+1}}\nabla\hat{Z}_r\mathrm{d}W_r+\mathcal{O}(\Delta t),\end{aligned}$$

and therefore

$$\begin{aligned}-Q_{t_n}\int_{t_n}^{t_{n+1}}\frac{\tilde{\rho}_s}{\sigma_s}\hat{Z}_s\mathrm{d}\overleftarrow{V}_s=&-Q_{t_n}\hat{Y}'_{t_{n+1}}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s\mathrm{d}\overleftarrow{V}_s\\&-Q_{t_n}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s\int_s^{t_{n+1}}\left(\hat{h}'_r\hat{Y}_r\nabla\hat{X}_r+\hat{h}_r\nabla\hat{Y}_r+\frac{\tilde{\rho}_r}{\sigma_r}\nabla\hat{Z}_r\right)\mathrm{d}\overleftarrow{V}_r\mathrm{d}\overleftarrow{V}_s\\&-Q_{t_n}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s\int_s^{t_{n+1}}\nabla\hat{Z}_r\mathrm{d}W_r\mathrm{d}\overleftarrow{V}_s+\mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right).\end{aligned}$$

Since W and V are two independent Brownian motions,

$$\begin{aligned}&\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s\int_s^{t_{n+1}}\nabla\hat{Z}_r\mathrm{d}W_r\mathrm{d}\overleftarrow{V}_s\right]\\=&\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s\int_s^{t_{n+1}}(\nabla\hat{Z}_r-\nabla\hat{Z}_{t_n})\mathrm{d}W_r\mathrm{d}\overleftarrow{V}_s\right]\le C(\Delta t)^{\frac{3}{2}}.\end{aligned} \tag{3.25}$$

As a result,

$$\begin{aligned}(\mathbf{iii-2})=&\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\frac{\tilde{\rho}_s}{\sigma_s}\hat{Z}_s\mathrm{d}\overleftarrow{V}_s\right]\\&-\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\tilde{\rho}_s\int_s^{t_{n+1}}\left(\hat{h}'_r\hat{Y}_r\nabla\hat{X}_r+\hat{h}_r\nabla\hat{Y}_r+\frac{\tilde{\rho}_r}{\sigma_r}\nabla\hat{Z}_r\right)\mathrm{d}\overleftarrow{V}_r\mathrm{d}\overleftarrow{V}_s\right]+\mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right)\\=&\mathbb{E}_0^x\left[Q_{t_n}\int_{t_n}^{t_{n+1}}\frac{\tilde{\rho}_s}{\sigma_s}\hat{Z}_s\mathrm{d}\overleftarrow{V}_s\right]-\mathbb{E}_0^x\left[\tilde{\lambda}_{t_n}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\mathrm{d}\overleftarrow{V}_r\mathrm{d}\overleftarrow{V}_s\right]+\mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right),\end{aligned} \tag{3.26}$$

where $\tilde{\lambda}_{t_n}=Q_{t_n}\tilde{\rho}_{t_{n+1}}\left(\hat{h}'_{t_n}\hat{Y}_{t_{n+1}}\nabla\hat{X}_{t_n}+\hat{h}_{t_n}\nabla\hat{Y}_{t_{n+1}}+\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}\nabla\hat{Z}_{t_{n+1}}\right)$ is independent of $\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}d\overleftarrow{V}_r d\overleftarrow{V}_s=\frac{1}{2}\left((V_{t_{n+1}}-V_{t_n})^2-\Delta t\right)$. By an argument similar to (3.16), we obtain

$$\sum_{n=0}^{N-1}\mathbb{E}_0^x\left[\tilde{\lambda}_{t_n}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\mathrm{d}\overleftarrow{V}_r\mathrm{d}\overleftarrow{V}_s\right]\overset{N\to\infty}{\longrightarrow}0\text{ in }\mathcal{L}^1(\Omega,\tilde{\mathbb{P}}_x). \tag{3.27}$$

Collecting estimates (3.21)–(3.24) and (3.26) into (3.19); then inserting (3.19), (3.17) and (3.18) into (3.9) we finally obtain

$$\Delta_n=\mathbb{E}_0^x\left[\int_{t_n}^{t_{n+1}}\hat{h}_sQ_s\mathrm{d}V_s\hat{Y}_{t_{n+1}}-Q_{t_n}\int_{t_n}^{t_{n+1}}\hat{h}_s\hat{Y}_s\mathrm{d}\overleftarrow{V}_s+\nu_n\right]+\mathcal{O}\left((\Delta t)^{\frac{3}{2}}\right)$$

$$:= \mathbb{E}_0^x[\alpha_n] + \mathbb{E}_0^x[\beta_n] + \mathbb{E}_0^x[\gamma_n] + \mathbb{E}_0^x[\nu_n] + \mathcal{O}\Big((\Delta t)^{\frac{3}{2}}\Big), \tag{3.28}$$

where $\{\nu_n\}$ is defined as in (3.15) and satisfies (3.16), and

$$\begin{aligned}
\alpha_n &:= \int_{t_n}^{t_{n+1}} \Big(Q_s \hat{h}_s - Q_{t_n} \hat{h}_{t_n}\Big) \hat{Y}_{t_{n+1}} \mathrm{d}V_s, \\
\beta_n &:= \int_{t_n}^{t_{n+1}} Q_{t_n} \Big(\hat{h}_{t_{n+1}} \hat{Y}_{t_{n+1}} - \hat{h}_s \hat{Y}_s\Big) \mathrm{d}\overleftarrow{V}_s, \\
\gamma_n &:= Q_{t_n} \hat{Y}_{t_{n+1}} \Big(\hat{h}_{t_n} - \hat{h}_{t_{n+1}}\Big) \cdot (V_{t_{n+1}} - V_{t_n}).
\end{aligned}$$

Part II.

In the second part of the proof, we show that $\sum_{n=0}^{N-1} \mathbb{E}_0^x[\alpha_n] \to 0$, $\sum_{n=0}^{N-1} \mathbb{E}_0^x[\beta_n] \to 0$, and $\sum_{n=0}^{N-1} \mathbb{E}_0^x[\gamma_n] \to 0$ in $\mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x)$ as $N \to \infty$.

First write $\alpha_n = \alpha_n^{(1)} + \alpha_n^{(2)} + \alpha_n^{(3)}$ with

$$\begin{aligned}
\alpha_n^{(1)} &:= \int_{t_n}^{t_{n+1}} (Q_s - Q_{t_n}) \hat{h}_s \mathrm{d}V_s \cdot \hat{Y}_{t_{n+1}}, \\
\alpha_n^{(2)} &:= \int_{t_n}^{t_{n+1}} Q_{t_n} (\hat{h}_s - \hat{h}_{t_n}) \mathrm{d}V_s \cdot \hat{Y}_{t_n}, \\
\alpha_n^{(3)} &:= \int_{t_n}^{t_{n+1}} Q_{t_n} (\hat{h}_s - \hat{h}_{t_n}) \mathrm{d}V_s \cdot \Big(\hat{Y}_{t_{n+1}} - \hat{Y}_{t_n}\Big).
\end{aligned}$$

Denote by $\tilde{\mathbb{E}}_x$ the expectation with respect to $\tilde{\mathbb{P}}_x$, where $\tilde{\mathbb{P}}_x := \tilde{\mathbb{P}}(\cdot|U_0 = x)$ is the induced probability measure. Noticing that $\hat{Y}_{t_n} = Y_{t_n}(U_{t_n})$ due to $\hat{X}_{t_n} = U_{t_n}$ given in (3.6), and that h_t is a bounded function, we apply Itô's formula to $(Q_s - Q_{t_n})$ in $\alpha_n^{(1)}$ to get

$$\begin{aligned}
\tilde{\mathbb{E}}_x\Bigg[\Bigg|\mathbb{E}_0^x\Big[\sum_{n=0}^{N-1} \alpha_n^{(1)}\Big]\Bigg|\Bigg] &= \tilde{\mathbb{E}}_x\Bigg[\Bigg|\mathbb{E}_0^x\Big[\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \hat{h}_r Q_r \mathrm{d}V_r \hat{h}_s \mathrm{d}V_s \cdot \hat{Y}_{t_{n+1}}\Big]\Bigg|\Bigg] \\
&\le \tilde{\mathbb{E}}_x\Bigg[\Bigg|\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \hat{h}_r (Q_r - Q_{t_n}) \mathrm{d}V_r \hat{h}_s \mathrm{d}V_s \cdot \hat{Y}_{t_{n+1}}\Bigg|\Bigg] \\
&\quad + \tilde{\mathbb{E}}_x\Bigg[\Bigg|\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \hat{h}_r Q_{t_n} \mathrm{d}V_r \hat{h}_s \mathrm{d}V_s \cdot \hat{Y}_{t_{n+1}}\Bigg|\Bigg] \\
&\le C \sum_{n=0}^{N-1} (\Delta t)^{\frac{3}{2}} + C \tilde{\mathbb{E}}_x\Bigg[\Bigg|\sum_{n=0}^{N-1} Q_{t_n} \hat{Y}_{t_{n+1}} \cdot \frac{1}{2}\Big((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\Big)\Bigg|\Bigg].
\end{aligned}$$

and from the fact that

$$\sum_{n=0}^{N-1} Q_{t_n} \hat{Y}_{t_{n+1}} \cdot \frac{1}{2}\big((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\big) \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x)$$

we have

$$\sum_{n=0}^{N-1} \mathbb{E}_0^x\big[\alpha_n^{(1)}\big] = \mathbb{E}_0^x\Big[\sum_{n=0}^{N-1} \alpha_n^{(1)}\Big] \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.29}$$

For $\alpha_n^{(2)}$, we apply Itô's formula to h_s to get

$$\begin{aligned}\mathbb{E}_0^x[\alpha_n^{(2)}] =& \mathbb{E}_0^x\Big[\int_{t_n}^{t_{n+1}} Q_{t_n}\cdot\Big(\int_{t_n}^{s}[\hat{\mu}_r\cdot\hat{h}'_r+\frac{(\sigma_r)^2}{2}\cdot\hat{h}''_r]\mathrm{d}r+\int_{t_n}^{s}\sigma_r\cdot\hat{h}'_r\mathrm{d}W_r\Big)\cdot\hat{Y}_{t_n}\mathrm{d}V_s\Big]\\ =& \mathbb{E}_0^x\Big[\int_{t_n}^{t_{n+1}} Q_{t_n}\hat{Y}_{t_n}\cdot\int_{t_n}^{s}[\hat{\mu}_r\cdot\hat{h}'_r+\frac{(\sigma_r)^2}{2}\hat{h}''_r]\mathrm{d}r\mathrm{d}V_s\Big].\end{aligned}$$

Since μ_t, σ_t, h' and h'' are all bounded, we have

$$\tilde{\mathbb{E}}_x[|\mathbb{E}_0^x[\alpha_n^{(2)}]|]\le C\Delta t^{3/2},$$

and consequently

$$\sum_{n=0}^{N-1}\mathbb{E}_0^x[\alpha_n^{(2)}]\to 0 \text{ in } \mathcal{L}^1(\Omega,\tilde{\mathbb{P}}_x). \tag{3.30}$$

Moreover, it follows from Hölder's inequality, Lemma 2.1 and Lemma 3.1 that

$$\begin{aligned}\tilde{\mathbb{E}}_x[|\mathbb{E}_0^x[\alpha_n^{(3)}]|] =& \tilde{\mathbb{E}}_x\Big[\Big|\mathbb{E}_0^x\Big[\int_{t_n}^{t_{n+1}} Q_{t_n}(\hat{h}_s-\hat{h}_{t_n})\mathrm{d}V_s\cdot(\hat{Y}_{t_{n+1}}-\hat{Y}_{t_n})\Big]\Big|\Big]\\ \le& \Big(\tilde{\mathbb{E}}_x\Big[\Big(\int_{t_n}^{t_{n+1}} Q_{t_n}(\hat{h}_s-\hat{h}_{t_n})\mathrm{d}V_s\Big)^2\Big]\Big)^{\frac{1}{2}}\cdot\Big(\tilde{\mathbb{E}}_x\big[(\hat{Y}_{t_{n+1}}-\hat{Y}_{t_n})^2\big]\Big)^{\frac{1}{2}}\\ \le& C\Delta t^{3/2},\end{aligned}$$

and hence

$$\sum_{n=0}^{N-1}\mathbb{E}_0^x[\alpha_n^{(3)}]\to 0 \text{ in } \mathcal{L}^1(\Omega,\tilde{\mathbb{P}}_x). \tag{3.31}$$

It follows immediately from (3.29), (3.30) and (3.31) that

$$\sum_{n=0}^{N-1}\mathbb{E}_0^x[\alpha_n]\to 0 \text{ in } \mathcal{L}^1(\Omega,\tilde{\mathbb{P}}_x). \tag{3.32}$$

For the term β_n in (3.28), we have

$$\begin{aligned}\beta_n =& \int_{t_n}^{t_{n+1}} Q_{t_n}\Big((\hat{h}_{t_{n+1}}-\hat{h}_s)\hat{Y}_{t_{n+1}}+\hat{h}_s(\hat{Y}_{t_{n+1}}-\hat{Y}_s)\Big)\mathrm{d}\overleftarrow{V}_s\\ =& \int_{t_n}^{t_{n+1}} Q_{t_n}(\hat{h}_{t_{n+1}}-\hat{h}_s)\hat{Y}_{t_n}\mathrm{d}\overleftarrow{V}_s+\int_{t_n}^{t_{n+1}} Q_{t_n}\hat{h}_s(\hat{Y}_{t_{n+1}}-\hat{Y}_s)\mathrm{d}\overleftarrow{V}_s\\ &+\int_{t_n}^{t_{n+1}} Q_{t_n}(\hat{h}_{t_{n+1}}-\hat{h}_s)\mathrm{d}\overleftarrow{V}_s\cdot(\hat{Y}_{t_{n+1}}-\hat{Y}_{t_n})\\ =& \beta_n^{(1)}+\beta_n^{(2)}+\beta_n^{(3)}\end{aligned}$$

with

$$\beta_n^{(1)}=\int_{t_n}^{t_{n+1}} Q_{t_n}(\hat{h}_{t_{n+1}}-\hat{h}_s)\hat{Y}_{t_n}\mathrm{d}\overleftarrow{V}_s,$$

$$\beta_n^{(2)} = \int_{t_n}^{t_{n+1}} Q_{t_n} \hat{h}_s (\hat{Y}_{t_{n+1}} - \hat{Y}_s) d\overleftarrow{V}_s,$$

$$\beta_n^{(3)} = \int_{t_n}^{t_{n+1}} Q_{t_n} (\hat{h}_{t_{n+1}} - \hat{h}_s) d\overleftarrow{V}_s \cdot (\hat{Y}_{t_{n+1}} - \hat{Y}_{t_n}).$$

Following similar approaches to estimate $\alpha_n^{(2)}$ and $\alpha_n^{(3)}$, we obtain

$$\tilde{\mathbb{E}}_x[|\mathbb{E}_0^x[\beta_n^{(1)}]|] \leq C\Delta t^{3/2}, \quad \tilde{\mathbb{E}}_x[|\mathbb{E}_0^x[\beta_n^{(3)}]|] \leq C\Delta t^{3/2},$$

and thus

$$\sum_{n=0}^{N-1} \mathbb{E}_0^x[\beta_n^{(1)}] \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x), \quad \sum_{n=0}^{N-1} \mathbb{E}_0^x[\beta_n^{(3)}] \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.33}$$

From the BDSDE in (3.1), Proposition 2.1, Lemma 2.1, and (3.25), we get

$$\begin{aligned}
\sum_{n=0}^{N-1} \mathbb{E}_0^x[\beta_n^{(2)}] &= \sum_{n=0}^{N-1} \mathbb{E}_0^x\Big[\int_{t_n}^{t_{n+1}} Q_{t_n} \hat{h}_s \Big(\int_s^{t_{n+1}} \hat{Z}_r dW_r - \int_s^{t_{n+1}} (\hat{h}_r \hat{Y}_r + \frac{\tilde{\rho}_r}{\sigma_r} \hat{Z}_r) d\overleftarrow{V}_r\Big) d\overleftarrow{V}_s\Big] \\
&= \sum_{n=0}^{N-1} \mathbb{E}_0^x\Big[\int_{t_n}^{t_{n+1}} Q_{t_n} \hat{h}_s \big(-\int_s^{t_{n+1}} (\hat{h}_r \hat{Y}_{t_{n+1}} + \frac{\tilde{\rho}_r}{\sigma_r} \hat{Z}_{t_{n+1}}) d\overleftarrow{V}_r\big) d\overleftarrow{V}_s\Big] + \mathcal{O}((\Delta t)^{\frac{3}{2}}) \\
&\quad + \sum_{n=0}^{N-1} \mathbb{E}_0^x\Big[\int_{t_n}^{t_{n+1}} Q_{t_n} \hat{h}_s \big(\int_s^{t_{n+1}} \hat{h}_r (\hat{Y}_{t_{n+1}} - \hat{Y}_r) + \frac{\tilde{\rho}_r}{\sigma_r} (\hat{Z}_{t_{n+1}} - \hat{Z}_r) d\overleftarrow{V}_r\big) d\overleftarrow{V}_s\Big].
\end{aligned}$$

Taking conditional expectation $\tilde{\mathbb{E}}_x$ of the absolute value of the above equation and using the facts

$$\tilde{\mathbb{E}}_x\big[(\hat{Y}_{t_{n+1}} - \hat{Y}_r)^2\big] \leq C(\Delta t)^3, \quad \tilde{\mathbb{E}}_x\big[(\hat{Z}_{t_{n+1}} - \hat{Z}_r)^2\big] \leq C(\Delta t)^3,$$

$$\begin{aligned}
&\tilde{\mathbb{E}}_x\Big[\Big|\mathbb{E}_0^x\big[\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} Q_{t_n} \hat{h}_s \big(-\int_s^{t_{n+1}} \hat{h}_r \hat{Y}_{t_{n+1}} d\overleftarrow{V}_r\big) d\overleftarrow{V}_s\big]\Big|\Big] \\
\leq & C\tilde{\mathbb{E}}_x\Big[\Big|\sum_{n=0}^{N-1} Q_{t_n} \hat{Y}_{t_{n+1}} \cdot \frac{1}{2}\big((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\big)\Big|\Big],
\end{aligned}$$

and

$$\sum_{n=0}^{N-1} Q_{t_n} \hat{Y}_{t_{n+1}} \cdot \frac{1}{2}\big((V_{t_{n+1}} - V_{t_n})^2 - \Delta t\big) \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x)$$

we obtain

$$\sum_{n=0}^{N-1} \mathbb{E}_0^x[\beta_n^{(2)}] = \mathbb{E}_0^x\big[\sum_{n=0}^{N-1} \beta_n^{(2)}\big] \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.34}$$

The above relation (3.34), together with (3.33), gives

$$\sum_{n=0}^{N-1} \mathbb{E}_0^x[\beta_n] = \mathbb{E}_0^x\big[\sum_{n=0}^{N-1} \beta_n\big] \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.35}$$

It remains to estimate γ_n. Applying Itô's formula to h_t, it's easy to verify that

$$|\mathbb{E}_0^x[\hat{h}_{t_{n+1}} - \hat{h}_{t_n}]| \leq C\Delta t. \tag{3.36}$$

Since $\hat{h}_{t_{n+1}} - \hat{h}_{t_n}$ is independent of $Q_{t_n}\hat{Y}_{t_n}(V_{t_{n+1}} - V_{t_n})$,

$$\begin{aligned}
\mathbb{E}_0^x[\gamma_n] &= \mathbb{E}_0^x\big[Q_{t_n}\hat{Y}_{t_n}\big(\hat{h}_{t_n} - \hat{h}_{t_{n+1}}\big)\cdot(V_{t_{n+1}} - V_{t_n})\big] \\
&\quad + \mathbb{E}_0^x\big[Q_{t_n}\big(\hat{Y}_{t_{n+1}} - \hat{Y}_{t_n}\big)\cdot\big(\hat{h}_{t_n} - \hat{h}_{t_{n+1}}\big)\cdot(V_{t_{n+1}} - V_{t_n})\big] \\
&= \mathbb{E}_0^x\big[\hat{h}_{t_n} - \hat{h}_{t_{n+1}}\big]\cdot\mathbb{E}_0^x\big[Q_{t_n}\hat{Y}_{t_n}\cdot(V_{t_{n+1}} - V_{t_n})\big] \\
&\quad + \mathbb{E}_0^x\big[Q_{t_n}\big(\hat{Y}_{t_{n+1}} - \hat{Y}_{t_n}\big)\cdot\big(\hat{h}_{t_n} - \hat{h}_{t_{n+1}}\big)\cdot(V_{t_{n+1}} - V_{t_n})\big].
\end{aligned}$$

Then, from estimate (3.36), Lemma 2.1, Lemma 3.1, we get

$$\begin{aligned}
\tilde{\mathbb{E}}_x\big[|\mathbb{E}_0^x[\gamma_n]|\big] \leq\; & C\Delta t\cdot\tilde{\mathbb{E}}_x\big[\big|Q_{t_n}\hat{Y}_{t_n}\cdot(V_{t_{n+1}} - V_{t_n})\big|\big] \\
& + \tilde{\mathbb{E}}_x\big[\big|Q_{t_n}\big(\hat{Y}_{t_{n+1}} - \hat{Y}_{t_n}\big)\cdot\big(\hat{h}_{t_n} - \hat{h}_{t_{n+1}}\big)\cdot(V_{t_{n+1}} - V_{t_n})\big|\big] \\
\leq\; & C(\Delta t)^{\frac{3}{2}}
\end{aligned} \tag{3.37}$$

and therefore

$$\sum_{n=0}^{N-1}\mathbb{E}_0^x[\gamma_n] = \mathbb{E}_0^x[\sum_{n=0}^{N-1}\gamma_n] \to \text{ in } L^1(\Omega, \tilde{\mathbb{P}}_x). \tag{3.38}$$

Finally with convergence results in (3.32), (3.35) and (3.38), the expression of Δ_n in (3.28) gives us

$$\sum_{n=0}^{N-1}\Delta_n \to 0 \text{ in } \mathcal{L}^1(\Omega, \tilde{\mathbb{P}}_x)$$

as desired. The proof is complete. □

3.2. Adjoint FBDSDEs. In this subsection, we provide a different perspective to consider the FBDSDE system. Specifically, we introduce the following FBDSDE system, in which the "forward SDE" (2.7) goes backward and the "backward SDE" (2.8) goes forward

$$\begin{cases}
d\overleftarrow{X}_t = \mu_t(\overleftarrow{X}_t)dt - \sigma_t d\overleftarrow{W}_t, \quad 0 \leq t \leq \tau & \text{(SDE)} \\
d\overrightarrow{Y}_t = -\mu_t'(\overleftarrow{X}_t)\overrightarrow{Y}_t dt - \overrightarrow{Z}_t d\overleftarrow{W}_t + \left(h(\overleftarrow{X}_t)\overrightarrow{Y}_t - \dfrac{\tilde{\rho}_t}{\sigma_t}\overrightarrow{Z}_t\right)dV_t, \quad 0 \leq t \leq \tau & \text{(BDSDE)} \\
\overleftarrow{X}_\tau = x, \qquad \overrightarrow{Y}_0 = p_0(\overleftarrow{X}_0), &
\end{cases} \tag{3.39}$$

where $0 \leq \tau \leq T$, $\int_t^T \cdot d\overleftarrow{W}_s$ is a backward Itô Integral and $\int_t^T \cdot dV_s$ is a standard forward Itô integral. Write the solution to (3.39) as $(\overleftarrow{X}_t^{T,x}, \overrightarrow{Y}_t^{T,x}, \overrightarrow{Z}_t^{T,x})$. Then by inverting the time index in the standard FBDSDE system, $\overleftarrow{X}_t^{T,x}$ is a $\mathcal{F}_{t,T}^W$ adapted stochastic process and the solution $\left(\overrightarrow{Y}_t^{T,x}, \overrightarrow{Z}_t^{T,x}\right)$ of the BDSDE in (3.39) is adapted to $\mathcal{F}_{t,T}^W \vee \mathcal{F}_t^V$.

In most literatures about FBDSDEs, the side condition of the solution "Y" is given at the terminal time T and the propagation direction is from T to 0, which is similar to the Feynman-Kac-type FBDSDE system (3.1). Although the adjoint relations for SPDEs are well-known, not many discussions have occurred in relating an FBDSDE system to its adjoint time-inverse

FBDSDE system. In what follows, we show that the solution $\overrightarrow{Y}_t$, which is initialized at the side condition $t=0$, is the adjoint stochastic process of Y_t defined in the FBDSDE system (3.1), and we denote by $\langle\cdot,\cdot\rangle$ the standard inner product in $\mathcal{L}^2(\mathbb{R})$, i.e. for any functions $\phi,\psi:\mathbb{R}\to\mathbb{R}$, $\langle\phi,\psi\rangle:=\int_{\mathbb{R}}\phi(x)\psi(x)\mathrm{d}x$.

Before we present the adjoint theorem (Theorem 3.2) for FBDSDE systems, we first introduce some regularity properties for μ and σ.

ASSUMPTION 3.1. *For $0\le s\le t\le T$, functions μ and σ satisfy*

$$|\mu_t(x)-\mu_s(x)|+|\mu_t'(x)-\mu_s'(x)|\le C|t-s|,\qquad |\sigma_t-\sigma_s|\le C|t-s|,$$

where C is a given positive constant independent of μ, σ, s and t.

Lemma 3.2 can be proved by using repeatedly the variational form of BDSDEs [31].

LEMMA 3.2. *Assume that $\mu\in\mathcal{C}_b^4$, $\phi\in\mathcal{C}_b^3$, $h\in\mathcal{C}_b^3$ and every derivative of μ, ϕ and h has bounded support in $\mathbb{R}$. Then for each $m_1=0,1,2$ and $m_2=0,1,2,3$, $Y_t^{(m_1)}$, $\overrightarrow{Y}_t^{(m_2)}$ have bounded support and satisfy*

$$\int_{\mathbb{R}}\tilde{\mathbb{E}}\left[\sup_{0\le t\le T}\left|Y_t^{(m_1)}\right|^2\right]\mathrm{d}x<\infty \text{ and } \int_{\mathbb{R}}\tilde{\mathbb{E}}\left[\sup_{0\le t\le T}\left|\overrightarrow{Y}_t^{(m_2)}\right|^2\right]\mathrm{d}x<\infty. \tag{3.40}$$

THEOREM 3.2. *Assume that, in addition to Assumption 3.1, σ is uniformly bounded, $\mu\in\mathcal{C}_b^4$, $\phi\in\mathcal{C}_b^3$, $h\in\mathcal{C}_b^3$ and each derivative of μ, ϕ and h has bounded support in $\mathbb{R}$. Then the process $R_t:=\left\langle Y_t,\overrightarrow{Y}_t\right\rangle$, $t\in[0,T]$ is a constant for almost all trajectories.*

Similar to the notation used in Section 3.1, we denote $\overrightarrow{Y}_t(x):=\overrightarrow{Y}_t^{t,x}$ and $\overrightarrow{Z}_t(x):=\overrightarrow{Z}_t^{t,x}$. In addition, for any non-negative integer m and function $f=f(x)$, we write $f^{(m)}:=\frac{\partial^m f}{\partial x^m}$. It's worthy noting that a similar duality described in Theorem 3.2 also occurs in the FBSDE (forward backward stochastic differential equation) systems. Actually, if we ignore the measurement process V_t in the optimal filtering problem, i.e. let $V_t\equiv 0$, the optimal filtering-type Feynman-Kac formula would become the classic Feynman-Kac formula, and both FBDSDE systems (3.1) and (3.39) become FBSDE systems. Therefore, the following proof will lead to the adjoint relation between FBSDEs by simply eliminating all the V_t and $\overleftarrow{V}_t$ terms. When we put back the observation V_t in the optimal filtering problem, the coefficient of observation noise resembles the coefficient of the multiplicative noise in the Zakai equation.

Proof. **(Proof of Theorem 3.2.)** First notice that according to [31], R_t has continuous paths a.s.. Thus it suffices to show that $R_s=R_t$ a.s. for all $s,t\in[0,T]$. For $0\le s\le t\le T$, let $s=t_0<t_1<\cdots<t_N=t$ be a temporal partition with uniform stepsize $t_{n+1}-t_n=\frac{t-s}{N}=\Delta t$.

For simplification of notations, we denote

$$\Delta V_{t_n}:=V_{t_{n+1}}-V_{t_n},\qquad Y_n:=Y_{t_n},\qquad Z_n:=Z_{t_n},\qquad \overrightarrow{Y}_n:=\overrightarrow{Y}_{t_n},\qquad \overrightarrow{Z}_n:=\overrightarrow{Z}_{t_n}.$$

By Corollary 2.2 in [31], we have

$$\begin{aligned}
&Y_n(x)=Y_{t_n}^{t_n,x}, \quad \overrightarrow{Y}_n(x) =\overrightarrow{Y}_{t_n}^{t_n,x}, \quad Y_{n+1}(X_{t_{n+1}}^{t_n,x})=Y_{t_{n+1}}^{t_n,x}, \quad \overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) =\overrightarrow{Y}_{t_n}^{t_{n+1},x},\\
&Z_n(x)=Z_{t_n}^{t_n,x}, \quad \overrightarrow{Z}_n(x) =\overrightarrow{Z}_{t_n}^{t_n,x}, \quad Z_{n+1}(X_{t_{n+1}}^{t_n,x})=Z_{t_{n+1}}^{t_n,x}, \quad \overrightarrow{Z}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) =\overrightarrow{Z}_{t_n}^{t_{n+1},x}.
\end{aligned}$$

Denote conditional expectations

$$\mathbb{E}[\cdot]:=\tilde{\mathbb{E}}[\cdot|\mathcal{F}_T^V],\quad \mathbb{E}_x^n[\cdot]:=\tilde{\mathbb{E}}[\cdot|\mathcal{F}_T^V,X_{t_n}=x],\quad \overleftarrow{\mathbb{E}}_x^n[\cdot]:=\tilde{\mathbb{E}}[\cdot|\mathcal{F}_T^V,\overleftarrow{X}_{t_n}=x].$$

It follows from the definitions of $\mathbb{E}_x^n$ and $\overleftarrow{\mathbb{E}}_x^n$ that

$$\mathbb{E}_x^n[Y_n]=Y_n(x), \qquad \overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]=\overrightarrow{Y}_n(x).$$

Without loss of generality suppose that $\Delta t<s\wedge(T-t)$ and define

$$Y_N=\frac{1}{\Delta t}\int_t^{t+\Delta t}Y_r\mathrm{d}r, \qquad \overrightarrow{Y}_0=\frac{1}{\Delta t}\int_{s-\Delta t}^{s}\overrightarrow{Y}_r\mathrm{d}r.$$

For $n=0,1,\dots,N-1$, taking the conditional expectations $\mathbb{E}_x^n$ and $\overleftarrow{\mathbb{E}}_x^{n+1}$ of temporal discretized approximations of the BDSDEs in (3.1) and (3.39). Specifically, we use the Euler scheme to approximate both the deterministic and stochastic integrals under the above temporal partition (see [6]) and derive

$$\mathbb{E}_x^n[Y_n]=\mathbb{E}_x^n[Y_{n+1}]+\mathbb{E}_x^n[h_{n+1}Y_{n+1}]\Delta V_{t_n}+\mathbb{E}_x^n\left[\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}Z_{n+1}\right]\Delta V_{t_n}+R_n, \tag{3.41}$$

$$\begin{aligned}\overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_{n+1}]=&\overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n]+\overleftarrow{\mathbb{E}}_x^n\left[-\overleftarrow{\mu}'_n\overrightarrow{Y}_n\right]\Delta t\\ &+\overleftarrow{\mathbb{E}}_x^{n+1}\left[\overleftarrow{h}_n\overrightarrow{Y}_n\right]\Delta V_{t_n}-\overleftarrow{\mathbb{E}}_x^{n+1}\left[\frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}}\overrightarrow{Z}_n\right]\Delta V_{t_n}+\tilde{R}_n,\end{aligned} \tag{3.42}$$

where

$$h_{n+1}:=h(X_{t_{n+1}}), \qquad \overleftarrow{\mu}'_n:=b'_{t_n}(\overleftarrow{X}_{t_n}), \qquad \overleftarrow{h}_n:=h(\overleftarrow{X}_{t_n}),$$

and we denote R_n to be the truncation error from the Euler scheme approximation to Equation (3.1), where

$$R_n=\int_{t_n}^{t_{n+1}}\left(h(X_s)Y_s+\frac{\tilde{\rho}_s}{\sigma_s}Z_s\right)\mathrm{d}\overleftarrow{V}_s-\left(\mathbb{E}_x^n[h_{n+1}Y_{n+1}]\Delta V_{t_n}+\mathbb{E}_x^n\left[\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}Z_{n+1}\right]\Delta V_{t_n}\right).$$

Similarly, the truncation error $\tilde{R}_n$ in approximating (3.39) is defined by

$$\begin{aligned}\tilde{R}_n=&\int_{t_n}^{t_{n+1}}-\mu'_s(\overleftarrow{X}_s)\overrightarrow{Y}_s\mathrm{d}s+\int_{t_n}^{t_{n+1}}\left(h(\overleftarrow{X}_s)\overrightarrow{Y}_s-\frac{\tilde{\rho}_s}{\sigma_t}\overrightarrow{Z}_s\right)\mathrm{d}V_s\\ &+\overleftarrow{\mathbb{E}}_x^n\left[\overleftarrow{\mu}'_n\overrightarrow{Y}_n\right]\Delta t-\overleftarrow{\mathbb{E}}_x^{n+1}\left[\overleftarrow{h}_n\overrightarrow{Y}_n\right]\Delta V_{t_n}+\overleftarrow{\mathbb{E}}_x^{n+1}\left[\frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}}\overrightarrow{Z}_n\right]\Delta V_{t_n}.\end{aligned}$$

The convergence results for the above scheme can be found in [6].

By the definition of expectations $\mathbb{E}_x^n$ and $\overleftarrow{\mathbb{E}}_x^{n+1}$,

$$\mathbb{E}_x^n[h_{n+1}]=\mathbb{E}\left[h(X_{t_{n+1}}^{t_n,x})\right], \quad \overleftarrow{\mathbb{E}}_x^{n+1}[\overleftarrow{\mu}'_n]=\mathbb{E}\left[\mu'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\right], \quad \overleftarrow{\mathbb{E}}_x^{n+1}\left[\overleftarrow{h}_n\right]=\mathbb{E}\left[h(\overleftarrow{X}_{t_n}^{t_{n+1},x})\right].$$

Multiplying (3.41) by $\overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]$ and (3.42) by $\mathbb{E}_x^{n+1}[Y_{n+1}]$, then taking integral with respect to $\mathrm{d}x$, we obtain

$$\begin{aligned}\left\langle\mathbb{E}_x^n[Y_n],\overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\right\rangle=&\left\langle\mathbb{E}_x^n[Y_{n+1}],\overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\right\rangle+\left\langle\mathbb{E}_x^n[h_{n+1}Y_{n+1}],\overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\right\rangle\Delta V_{t_n}\\ &+\left\langle\mathbb{E}_x^n\left[\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}Z_{n+1}\right],\overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\right\rangle\Delta V_{t_n}\end{aligned} \tag{3.43}$$

and

$$\begin{aligned}&\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_{n+1}],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\\&=\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_{n}],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle+\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}\Big[-\overleftarrow{\mu}'_n\overrightarrow{Y}_n\Big],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\Delta t\\&\quad+\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}\Big[\overleftarrow{h}_n\overrightarrow{Y}_n\Big],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\Delta V_{t_n}-\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}\Big[\frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}}\overrightarrow{Z}_n\Big],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\Delta V_{t_n}.\end{aligned}\tag{3.44}$$

Subtraction of (3.44) from (3.43) results in

$$\begin{aligned}&\Big\langle \mathbb{E}_x^{n}[Y_n],\overleftarrow{\mathbb{E}}_x^{n}[\overrightarrow{Y}_n]\Big\rangle-\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_{n+1}],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\\&=\underbrace{\Big\langle \mathbb{E}_x^{n}[Y_{n+1}],\overleftarrow{\mathbb{E}}_x^{n}[\overrightarrow{Y}_n]-\overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n]\Big\rangle+\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n],\mathbb{E}_x^{n}[Y_{n+1}]-\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle}_{(\mathbf{iv})}\\&\quad+\underbrace{\Big\langle \mathbb{E}_x^{n}[h_{n+1}Y_{n+1}],\overleftarrow{\mathbb{E}}_x^{n}[\overrightarrow{Y}_n]\Big\rangle\Delta V_{t_n}-\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}\Big[\overleftarrow{h}_n\overrightarrow{Y}_n\Big],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\Delta V_{t_n}}_{(\mathbf{v})}\\&\quad+\underbrace{\Big\langle \mathbb{E}_x^{n}\Big[\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}Z_{n+1}\Big],\overleftarrow{\mathbb{E}}_x^{n}[\overrightarrow{Y}_n]\Big\rangle\Delta V_{t_n}+\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}\Big[\frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}}\overrightarrow{Z}_n\Big],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\Delta V_{t_n}}_{(\mathbf{vi})}\\&\quad-\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}\Big[-\overleftarrow{\mu}'_n\overrightarrow{Y}_n\Big],\mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle\Delta t.\end{aligned}\tag{3.45}$$

In what follows, we prove that by taking the sum of Equation (3.45) from $n=0$ to $n=N-1$, the right-hand side of the resulting equation converges to 0 as $\Delta t\to 0$, which will lead to the desired result of this theorem. To this end we estimate terms (**iv**), (**v**) and (**vi**) one by one.

(**iv**) By the definitions $\overleftarrow{\mathbb{E}}_x^{n}$ and $\mathbb{E}_x^{n}$, we have

$$\begin{aligned}\overleftarrow{\mathbb{E}}_x^{n}[\overrightarrow{Y}_n]-\overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n]&=\mathbb{E}\Big[\overrightarrow{Y}_n(x)-\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\Big],\\ \mathbb{E}_x^{n}[Y_{n+1}]-\mathbb{E}_x^{n+1}[Y_{n+1}]&=\mathbb{E}\Big[Y_{n+1}(X_{t_{n+1}}^{t_n,x})-Y_{n+1}(x)\Big].\end{aligned}$$

It follows from Itô's formula that

$$\begin{aligned}\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})=\overrightarrow{Y}_n(x)&+\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_s^{t_{n+1},x})+\frac{(\sigma_s)^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s\\&+\int_{t_n}^{t_{n+1}}\sigma_s\overrightarrow{Y}'_n(\overleftarrow{X}_s^{t_{n+1},x})\mathrm{d}\overleftarrow{W}_s,\end{aligned}\tag{3.46}$$

$$\begin{aligned}Y_{n+1}(X_{t_{n+1}}^{t_n,x})=\ Y_{n+1}(x)&+\int_{t_n}^{t_{n+1}}\big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x})+\frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s\\&+\int_{t_n}^{t_{n+1}}\sigma_sY'_{n+1}(X_s^{t_n,x})\mathrm{d}W_s.\end{aligned}\tag{3.47}$$

Taking conditional expectation $\mathbb{E}$ to Equations (3.46) and (3.47), we obtain

$$\overleftarrow{\mathbb{E}}_x^{n}[\overrightarrow{Y}_n]-\overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n]=-\mathbb{E}\Big[\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_s^{t_{n+1},x})+\frac{(\sigma_s)^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s\Big]$$

$$= -\mathbb{E}\Big[-\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) + \frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\Big]\cdot\Delta t + \overrightarrow{R}_n,$$

$$\mathbb{E}_x^n[Y_{n+1}] - \mathbb{E}_x^{n+1}[Y_{n+1}] = \mathbb{E}\Big[\int_{t_n}^{t_{n+1}}\big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x}) + \frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s\Big]$$
$$= \mathbb{E}\Big[\mu_n(X_{t_n}^{t_n,x})Y'_{n+1}(X_{t_n}^{t_n,x}) + \frac{(\sigma_{t_n})^2}{2}Y''_{n+1}(X_{t_n}^{t_n,x})\Big]\cdot\Delta t + R_n,$$

where

$$\overrightarrow{R}_n := -\mathbb{E}\Big[\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s\Big]$$
$$+\mathbb{E}\Big[-\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) + \frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\Big]\cdot\Delta t,$$

$$R_n := \mathbb{E}\Big[\int_{t_n}^{t_{n+1}}\big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x}) + \frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s\Big]$$
$$-\mathbb{E}\big[b_n(X_{t_n}^{t_n,x})Y'_{n+1}(X_{t_n}^{t_n,x}) + \frac{(\sigma_{t_n})^2}{2}Y''_{n+1}(X_{t_n}^{t_n,x})\big]\cdot\Delta t.$$

As a consequence

$$\Big\langle\mathbb{E}_x^n[Y_{n+1}],\overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n] - \overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n]\Big\rangle$$
$$= -\int_{\mathbb{R}}\mathbb{E}\Big[Y_{n+1}(X_{t_{n+1}}^{t_n,x})\Big]\Big(\mathbb{E}\big[-\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) + \frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big]\cdot\Delta t - \overrightarrow{R}_n\Big)\mathrm{d}x. \tag{3.48}$$

Similarly

$$\Big\langle\overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_n],\mathbb{E}_x^n[Y_{n+1}] - \mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle$$
$$= \int_{\mathbb{R}}\mathbb{E}\Big[\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\Big]\Big(\mathbb{E}\big[\mu_n(X_{t_n}^{t_n,x})Y'_{n+1}(X_{t_n}^{t_n,x}) + \frac{(\sigma_{t_n})^2}{2}Y''_{n+1}(X_{t_n}^{t_n,x})\big]\cdot\Delta t + R_n\Big)\mathrm{d}x. \tag{3.49}$$

Adding (3.48) to (3.49) we have that

$$\begin{aligned}(\mathbf{iv}) = \Big(&\underbrace{-\int_{\mathbb{R}}\mathbb{E}\Big[Y_{n+1}(X_{t_{n+1}}^{t_n,x})\Big]\mathbb{E}\big[-\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big]\mathrm{d}x}_{(\mathbf{iv-1})}\\
&+\underbrace{\int_{\mathbb{R}}\mathbb{E}\Big[\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\Big]\mathbb{E}\big[\mu_n(X_{t_n}^{t_n,x})Y'_{n+1}(X_{t_n}^{t_n,x})\big]\mathrm{d}x}_{(\mathbf{iv-2})}\Big)\cdot\Delta t\\
&+\Big(\underbrace{-\int_{\mathbb{R}}\mathbb{E}\Big[Y_{n+1}(X_{t_{n+1}}^{t_n,x})\Big]\mathbb{E}\big[\frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big]\mathrm{d}x}_{(\mathbf{iv-3})}\\
&+\underbrace{\int_{\mathbb{R}}\mathbb{E}\Big[\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\Big]\mathbb{E}\big[\frac{(\sigma_{t_n})^2}{2}Y''_{n+1}(X_{t_n}^{t_n,x})\big]\mathrm{d}x}_{(\mathbf{iv-4})}\Big)\cdot\Delta t + R_n^x,\end{aligned} \tag{3.50}$$

where $R_n^x = \int_{\mathbb{R}} \mathbb{E}[Y_{n+1}(X_{t_{n+1}}^{t_n,x})]\overrightarrow{R}_n dx + \int_{\mathbb{R}} \mathbb{E}[\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})]R_n \mathrm{d}x$. Again by using the Itô's formula we obtain

$$\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) = \overrightarrow{Y}'_n(x) + \int_{t_n}^{t_{n+1}} \big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\overrightarrow{Y}_n^{(3)}(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} \sigma_s \overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\mathrm{d}\overleftarrow{W}_s,$$
$$\overrightarrow{Y}''_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) = \overrightarrow{Y}''_n(x) + \int_{t_n}^{t_{n+1}} \big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}_n^{(3)}(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\overrightarrow{Y}_n^{(4)}(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} \sigma_s \overrightarrow{Y}_n^{(3)}(\overleftarrow{X}_s^{t_{n+1},x})\mathrm{d}\overleftarrow{W}_s,$$
$$\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) = \mu_n(x) + \int_{t_n}^{t_{n+1}} -\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\mu'_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\mu''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} \sigma_s \mu'_n(\overleftarrow{X}_s^{t_{n+1},x})\mathrm{d}\overleftarrow{W}_s.$$

Hence, the term $\mathbb{E}\big[\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big]$ on the right-hand side of (3.50) can be written as $\mathbb{E}\big[\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big] = \mu_n(x)\overrightarrow{Y}'_n(x) + \Gamma_n(x)$ with

$$\Gamma_n(x) = \mathbb{E}\Big[\mu_n(x)\cdot \int_{t_n}^{t_{n+1}} \big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\overrightarrow{Y}_n^{(3)}(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$+ \overrightarrow{Y}'_n(x)\cdot \int_{t_n}^{t_{n+1}} \big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\mu'_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\mu''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} \big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\mu'_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\mu''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$\cdot \int_{t_n}^{t_{n+1}} \big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}\overrightarrow{Y}_n^{(3)}(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} (\sigma_s)^2 \mu'_n(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\mathrm{d}s\Big].$$

As a result, the terms on the right-hand side of (3.50) can be rewritten as

$$(\mathbf{iv-1}) = \int_{\mathbb{R}} \big(Y_{n+1}(x)\mu_n(x)\overrightarrow{Y}'_n(x)\big)\mathrm{d}x + H_n^1, \tag{3.51}$$
$$(\mathbf{iv-2}) = \int_{\mathbb{R}} \big(\overrightarrow{Y}_n(x)\mu_n(x)Y'_{n+1}(x)\big)\mathrm{d}x + H_n^2, \tag{3.52}$$
$$(\mathbf{iv-3}) = -\int_{\mathbb{R}} \frac{(\sigma_{t_n})^2}{2} Y_{n+1}(x)\overrightarrow{Y}''_n(x)\mathrm{d}x + H_n^3, \tag{3.53}$$
$$(\mathbf{iv-4}) = \int_{\mathbb{R}} \frac{(\sigma_{t_n})^2}{2} \overrightarrow{Y}_n(x)Y''_{n+1}(x)\mathrm{d}x + H_n^4, \tag{3.54}$$

where

$$H_n^1 = \int_{\mathbb{R}} \Big\{ \Big(Y_{n+1}(x) + \mathbb{E}[\int_{t_n}^{t_{n+1}} \big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x}) + \frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s)]\Big)\cdot \Gamma_n(x)$$

$$+\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x})+\frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s\big)]\cdot\mu_n(x)\overrightarrow{Y}'_n(x)\Big\}\mathrm{d}x,$$

$$H_n^2=\int_{\mathbb{R}}\Big\{\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_s^{t_{n+1},x})+\frac{(\sigma_s)^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s]\cdot\mu_n(x)Y'_{n+1}(x)\Big\}\mathrm{d}x,$$

$$H_n^3=-\int_{\mathbb{R}}\Big\{Y_{n+1}(x)\cdot\frac{(\sigma_{t_n})^2}{2}\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}^{(3)}_n(\overleftarrow{X}_s^{t_{n+1},x})+\frac{(\sigma_s)^2}{2}\overrightarrow{Y}^{(4)}_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s]$$

$$+\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x})+\frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s\big)]\cdot\frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}''_n(x)$$

$$+\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(\mu_s(X_s^{t_n,x})Y'_{n+1}(X_s^{t_n,x})+\frac{(\sigma_s)^2}{2}Y''_{n+1}(X_s^{t_n,x})\big)\mathrm{d}s\big)]$$

$$\cdot\frac{(\sigma_{t_n})^2}{2}\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}^{(3)}_n(\overleftarrow{X}_s^{t_{n+1},x})+\frac{(\sigma_s)^2}{2}\overrightarrow{Y}^{(4)}_n(\overleftarrow{X}_s^{t_{n+1},x})\big)\mathrm{d}s]\Big\}\mathrm{d}x,$$

$$H_n^4=\int_{\mathbb{R}}\Big\{\mathbb{E}[\int_{t_n}^{t_{n+1}}\big(-\mu_s(\overleftarrow{X}_s^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_s^{t_{n+1},x})+\frac{(\sigma_s)^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_s^{t_{n+1},x})\big)]\mathrm{d}s\cdot\frac{(\sigma_{t_n})^2}{2}Y''_{n+1}(x)\Big\}\mathrm{d}x.$$

Integrating (3.51), (3.53) and (3.54) by parts, we obtain

$$(\mathbf{iv-1})=-\int_{\mathbb{R}}Y'_{n+1}(x)\mu_n(x)\overrightarrow{Y}_n(x)\mathrm{d}x-\int_{\mathbb{R}}Y_{n+1}(x)\mu'_n(x)\overrightarrow{Y}_n(x)\mathrm{d}x, \tag{3.55}$$

$$(\mathbf{iv-3})=\int_{\mathbb{R}}\frac{(\sigma_{t_n})^2}{2}Y'_{n+1}(x)\overrightarrow{Y}'_n(x)\mathrm{d}x, \tag{3.56}$$

$$(\mathbf{iv-4})=-\int_{\mathbb{R}}\frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}'_n(x)Y'_{n+1}(x)\mathrm{d}x. \tag{3.57}$$

Adding (3.51) to (3.52) and applying (3.55), gives

$$\begin{aligned}(\mathbf{iv-1})+(\mathbf{iv-2})=&-\int_{\mathbb{R}}\mathbb{E}[Y_{n+1}(X_{t_{n+1}}^{t_n,x})]\mathbb{E}\big[-\mu_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}'_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big]\mathrm{d}x\\&+\int_{\mathbb{R}}\mathbb{E}[\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})]\mathbb{E}\big[\mu_n(X_{t_n}^{t_n,x})Y'_{n+1}(X_{t_n}^{t_n,x})\big]\mathrm{d}x\\=&-\int_{\mathbb{R}}Y_{n+1}(x)\mu'_n(x)\overrightarrow{Y}_n(x)\mathrm{d}x+H_n^1+H_n^2.\end{aligned} \tag{3.58}$$

Similarly, adding (3.53) to (3.54) and applying (3.56) and (3.57) yields

$$\begin{aligned}(\mathbf{iv-3})+(\mathbf{iv-4})=&-\int_{\mathbb{R}}\mathbb{E}[Y_{n+1}(\tilde{X}_{t_{n+1}}^{t_n,x})]\mathbb{E}\big[\frac{(\sigma_{t_n})^2}{2}\overrightarrow{Y}''_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big]\mathrm{d}x\\&+\int_{\mathbb{R}}\mathbb{E}[\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})]\mathbb{E}\big[\frac{(\sigma_{t_n})^2}{2}Y''_{n+1}(\tilde{X}_{t_n}^{t_n,x})\big]\mathrm{d}x\\=&\,H_n^3+H_n^4.\end{aligned} \tag{3.59}$$

In summary, inserting (3.58) and (3.59) into (3.50) gives

$$(\mathbf{iv})=\big(-\int_{\mathbb{R}}Y_{n+1}(x)\mu'_n(x)\overrightarrow{Y}_n(x)\mathrm{d}x\big)\Delta t+(H_n^1+H_n^2+H_n^3+H_n^4)\Delta t+R_n^x. \tag{3.60}$$

We next estimate the term $(\mathbf{v})$ in (3.45). From the definition of $\mathbb{E}_x^n$ and $\overleftarrow{\mathbb{E}}_x^{n+1}$, one has

$$(\mathbf{v})=\int_{\mathbb{R}}\mathbb{E}[h(X_{t_{n+1}}^{t_n,x})Y_{n+1}(X_{t_{n+1}}^{t_n,x})\overrightarrow{Y}_n(x)]-\mathbb{E}[h(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})Y_{n+1}(x)]\mathrm{d}x\Delta V_{t_n}. \tag{3.61}$$

Applying Itô's formula to h on time interval $[t_n, t_{n+1}]$ gives

$$h(X_{t_{n+1}}^{t_n,x}) = h(x) + \int_{t_n}^{t_{n+1}} \mu_s(X_s^{t_n,x})h'(X_s^{t_n,x}) + \frac{(\sigma_s)^2}{2}h''(X_s^{t_n,x}))\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} \sigma_s h'(X_s^{t_n,x})\mathrm{d}W_s;$$
$$h(\overleftarrow{X}_{t_n}^{t_{n+1},x}) = h(x) + \int_{t_n}^{t_{n+1}} -\mu_s(\overleftarrow{X}_s^{t_{n+1},x})h'(\overleftarrow{X}_s^{t_{n+1},x}) + \frac{(\sigma_s)^2}{2}h''(\overleftarrow{X}_s^{t_{n+1},x}))\mathrm{d}s$$
$$+ \int_{t_n}^{t_{n+1}} \sigma_s h'(\overleftarrow{X}_s^{t_{n+1},x})\mathrm{d}\overleftarrow{W}_s,$$

and thus

$$\begin{aligned}
&\mathbb{E}[h(X_{t_{n+1}}^{t_n,x})Y_{n+1}(X_{t_{n+1}}^{t_n,x})\overrightarrow{Y}_n(x)] \\
&= h(x)Y_{n+1}(x)\overrightarrow{Y}_n(x) + \overrightarrow{Y}_n(x)\cdot\mathbb{E}\big[\big(h(X_{t_{n+1}}^{t_n,x}) - h(x)\big)Y_{n+1}(x) \\
&\quad + h(x)\big(Y_{n+1}(X_{t_{n+1}}^{t_n,x}) - Y_{n+1}(x)\big) + \big(h(X_{t_{n+1}}^{t_n,x}) - h(x)\big)\big(Y_{n+1}(X_{t_{n+1}}^{t_n,x}) - Y_{n+1}(x)\big)\big],
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}[h(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})Y_{n+1}(x)] \\
&= h(x)\overrightarrow{Y}_n(x)Y_{n+1}(x) + Y_{n+1}(x)\cdot\mathbb{E}\big[\big(h(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - h(x)\big)\overrightarrow{Y}_n(x) \\
&\quad + h(x)\big(\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - \overrightarrow{Y}_n(x)\big) + \big(h(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - h(x)\big)\big(\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - \overrightarrow{Y}_n(x)\big)\big].
\end{aligned}$$

Consequently (3.61) becomes

$$(\mathbf{v}) = G_n^1 \Delta V_{t_n}, \tag{3.62}$$

where

$$\begin{aligned}
G_n^1 = \int_{\mathbb{R}} \big\{&\overrightarrow{Y}_n(x)\cdot\mathbb{E}\big[\big(h(X_{t_{n+1}}^{t_n,x}) - h(x)\big)Y_{n+1}(x) \\
&\quad + h(x)\big(Y_{n+1}(X_{t_{n+1}}^{t_n,x}) - Y_{n+1}(x)\big) + \big(h(X_{t_{n+1}}^{t_n,x}) - h(x)\big)\big(Y_{n+1}(X_{t_{n+1}}^{t_n,x}) - Y_{n+1}(x)\big)\big] \\
&- Y_{n+1}(x)\cdot\mathbb{E}\big[\big(h(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - h(x)\big)\overrightarrow{Y}_n(x) \\
&\quad + h(x)\big(\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - \overrightarrow{Y}_n(x)\big) + \big(h(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - h(x)\big)\big(\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) - \overrightarrow{Y}_n(x)\big)\big]\big\}\mathrm{d}x.
\end{aligned}$$

Finally, we consider the term $(\mathbf{vi})$ in (3.45). From the relation between Z_t and $\frac{\partial Y_t}{\partial x}$ given in (2.10), we know that

$$Z_{n+1}(X_{t_{n+1}}^{t_n,x}) = \frac{\partial Y_{n+1}(X_{t_{n+1}}^{t_n,x})}{\partial x}(\nabla X_{t_{n+1}}^{t_n,x})^{-1}\sigma_{t_{n+1}};$$
$$\overrightarrow{Z}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) = \frac{\partial \overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})}{\partial x}(\nabla \overleftarrow{X}_{t_n}^{t_{n+1},x})^{-1}\sigma_{t_n}.$$

Therefore we have

$$\Big\langle \mathbb{E}_x^n[\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}Z_{n+1}], \overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\Big\rangle = \mathbb{E}\Big[\int_{\mathbb{R}} \frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}}Z_{n+1}(X_{t_{n+1}}^{t_n,x})\cdot\overrightarrow{Y}_n(x)\mathrm{d}x\Big]$$

$$= \mathbb{E}\Big[\int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial Y_{n+1}(X_{t_{n+1}}^{t_n,x})}{\partial x} \cdot \overrightarrow{Y}_n(x) \cdot (\nabla X_{t_{n+1}}^{t_n,x})^{-1} \mathrm{d}x\Big]; \qquad (3.63)$$

$$\Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}[\frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}} \overrightarrow{Z}_n], \mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle = \mathbb{E}\Big[\int_{\mathbb{R}} \frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}} \overrightarrow{Z}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x}) \cdot Y_{n+1}(x)\mathrm{d}x\Big]$$
$$= \mathbb{E}\Big[\int_{\mathbb{R}} \tilde{\rho}_{t_n} \frac{\partial \overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})}{\partial x} \cdot Y_{n+1}(x) \cdot (\nabla \overleftarrow{X}_{t_n}^{t_{n+1},x})^{-1} \mathrm{d}x\Big], \qquad (3.64)$$

Adding (3.63) and (3.64) together, we obtain

$$(\mathbf{vi}) = \left(\mathbb{E}\Big[\int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial Y_{n+1}}{\partial x}(x) \cdot \overrightarrow{Y}_n(x)dx + \int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial \overrightarrow{Y}_n}{\partial x}(x) \cdot Y_{n+1}(x)dx\Big] + G_n^2\right) \Delta V_{t_n}, \qquad (3.65)$$

where

$$G_n^2 = \mathbb{E}\Big[\int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial Y_{n+1}}{\partial x}(X_{t_{n+1}}^{t_n,x}) \cdot \overrightarrow{Y}_n(x) \cdot (\nabla X_{t_{n+1}}^{t_n,x})^{-1}\mathrm{d}x - \int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial Y_{n+1}}{\partial x}(x) \cdot \overrightarrow{Y}_{t_n}(x)\mathrm{d}x\Big]$$
$$+ \mathbb{E}\Big[\int_{\mathbb{R}} \tilde{\rho}_{t_n} \frac{\partial \overrightarrow{Y}_n}{\partial x}(\overleftarrow{X}_{t_n}^{t_{n+1},x})) \cdot Y_{n+1}(x) \cdot (\nabla \overleftarrow{X}_{t_n}^{t_{n+1},x})^{-1}\mathrm{d}x - \int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial \overrightarrow{Y}_n}{\partial x}(x) \cdot Y_{n+1}(x)\mathrm{d}x\Big].$$

Integration by parts gives

$$\int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial Y_{n+1}}{\partial x}(x) \cdot \overrightarrow{Y}_n(x)\mathrm{d}x = -\int_{\mathbb{R}} \tilde{\rho}_{t_{n+1}} \frac{\partial \overrightarrow{Y}_n}{\partial x}(x) \cdot Y_{n+1}(x)\mathrm{d}x,$$

and therefore

$$G_n^2 = \Big\langle \mathbb{E}_x^n[\frac{\tilde{\rho}_{t_{n+1}}}{\sigma_{t_{n+1}}} Z_{n+1}], \overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\Big\rangle + \Big\langle \overleftarrow{\mathbb{E}}_x^{n+1}[\frac{\tilde{\rho}_{t_n}}{\sigma_{t_n}} \overrightarrow{Z}_n], \mathbb{E}_x^{n+1}[Y_{n+1}]\Big\rangle. \qquad (3.66)$$

Collecting (3.60), (3.62) and (3.66), and putting into Equation (3.45) results in

$$\Big\langle \mathbb{E}_x^n[Y_n], \overleftarrow{\mathbb{E}}_x^n[\overrightarrow{Y}_n]\Big\rangle - \Big\langle \mathbb{E}_x^{n+1}[Y_{n+1}], \overleftarrow{\mathbb{E}}_x^{n+1}[\overrightarrow{Y}_{n+1}]\Big\rangle$$
$$= (-\int_{\mathbb{R}} Y_{n+1}(x)\mu_n'(x)\overrightarrow{Y}_n(x)\mathrm{d}x)\Delta t + (H_n^1 + H_n^2 + H_n^3 + H_n^4)\Delta t + R_n^x$$
$$+ \int_{\mathbb{R}} \mathbb{E}[\mu_n'(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})]Y_{n+1}(x)\mathrm{d}x\Delta t + G_n \Delta V_{t_n}$$
$$= H_n\Delta t + R_n^x + G_n\Delta V_{t_n} + F_n\Delta t, \qquad (3.67)$$

where

$$H_n = H_n^1 + H_n^2 + H_n^3 + H_n^4, \quad G_n = G_n^1 + G_n^2$$

and

$$F_n = \int_{\mathbb{R}} Y_{n+1}(x)\Big(\mathbb{E}\big[\mu_n'(\overleftarrow{X}_{t_n}^{t_{n+1},x})\overrightarrow{Y}_n(\overleftarrow{X}_{t_n}^{t_{n+1},x})\big] - \mu_n'(x)\overrightarrow{Y}_n(x)\Big)\mathrm{d}x.$$

Sum (3.67) from $n=0$ to $n=N-1$ to get

$$\Big\langle \mathbb{E}_x^0[Y_0], \overleftarrow{\mathbb{E}}_x^0[\overrightarrow{Y}_0]\Big\rangle - \Big\langle \overleftarrow{\mathbb{E}}_x^N[\overrightarrow{Y}_N], \mathbb{E}_x^N[Y_N]\Big\rangle = \sum_{n=0}^{N-1}(H_n\Delta t + R_n^x + G_n\Delta V_{t_n} + F_n\Delta t). \qquad (3.68)$$

From definitions of H_n , R_n^x, G_n and F_n, it's straightforward to verify that $\mathbb{E}[(H_n)^2]\leq C(\Delta t)^2$, $\mathbb{E}[(R_n^x)^2]\leq C(\Delta t)^4$,$\mathbb{E}[(G_n)^2]\leq C(\Delta t)^2$ and $\mathbb{E}[(F_n)^2]\leq C(\Delta t)$. Therefore,

$$\lim_{\Delta t\to 0}\sum_{n=0}^{N-1}(H_n\Delta t+R_n^x+G_n\Delta V_{t_n}+F_n\Delta t)=0,\ a.s..$$

Also, since $\lim_{\Delta t\to 0}Y_0=Y_s$ and $\lim_{\Delta t\to 0}Y_N=Y_t$, we have

$$\left\langle Y_s,\overrightarrow{Y}_s\right\rangle=\left\langle Y_t,\overrightarrow{Y}_t\right\rangle$$

as required. The proof is complete. ∎

3.3. FBDSDEs and the optimal filtering problem. Now we are ready to show that the solution $\overrightarrow{Y}$ of (3.39) solves the optimal filter problem.

THEOREM 3.3. *Assume that the assumptions in Theorem 3.1 and Theorem 3.2 hold. Then*

$$\left\langle\overrightarrow{Y}_T,\phi\right\rangle=\tilde{\mathbb{E}}\left[\phi(U_T)Q_T\middle|\mathcal{F}_T^V\right],\qquad\forall\phi\in\mathcal{L}^\infty(\mathbb{R}^d).$$

Proof. By Theorem 3.2, one has

$$\left\langle\overrightarrow{Y}_T,Y_T\right\rangle=\left\langle\overrightarrow{Y}_0,Y_0\right\rangle.$$

Since $Y_T=\phi$ as given in (3.1), $\overrightarrow{Y}_0=p_0$ as given in (3.39) and $Y_0=\tilde{\mathbb{E}}_x\left[\phi(U_T)Q_T\middle|\mathcal{F}_T^V\right]$ as proved in Theorem 3.1, we have

$$\left\langle\overrightarrow{Y}_T,\phi\right\rangle=\int_{\mathbb{R}}p_0(x)\tilde{\mathbb{E}}_x[\phi(U_T)Q_T|\mathcal{F}_T^V]\mathrm{d}x.$$

Let φ be any bounded $\mathcal{F}_T^V$ measurable random variable,

$$\tilde{\mathbb{E}}_x\left[\left\langle\overrightarrow{Y}_T,\phi\right\rangle\varphi\right]=\int_{\mathbb{R}}p_0(x)\tilde{\mathbb{E}}_x[\phi(U_T)Q_T\varphi]\mathrm{d}x.$$

It then follows from the fact that $\tilde{\mathbb{P}}_x(\cdot|\mathcal{F}_T^V)=\tilde{\mathbb{P}}(\cdot|\mathcal{F}_T^V)$, and definition of $\tilde{\mathbb{P}}$

$$\tilde{\mathbb{E}}\left[\left\langle\overrightarrow{Y}_T,\phi\right\rangle\varphi\right]=\tilde{\mathbb{E}}[\phi(U_T)Q_T\varphi],$$

which completes the proof. ∎

REMARK 3.3. From (2.5), we can see that

$$\mathbb{E}\left[\phi(U_T)\middle|\mathcal{F}_T^V\right]=\frac{\left\langle\overrightarrow{Y}_T,Y_T\right\rangle}{\tilde{\mathbb{E}}\left[Q_t\middle|\mathcal{F}_t^V\right]}$$

Thus the solution $\overrightarrow{Y}_T$ of the FBDSDE (3.39) indeed provides an unnormalized filtering density for the optimal filtering problem.

4. Closing remarks

In this paper, a Feynmann-Kac-type BDSDE formula for optimal filter problems and its adjoint were derived. It was shown that the adjoint provides an unnormalized solution for the optimal filter problem (BSDE filter). As our preliminary work has shown, the BSDE filter has the potential to solve the optimal filter problem with more accuracy and less complexity than traditional filter methods.

REFERENCES

[1] C. Andrieu, A. Doucet, and R. Holenstein, *Particle Markov chain Monte Carlo methods*, J.R. Statist. Soc. B, 72(3):269–342, 2010. 1

[2] F. Bao, Y. Cao, and X. Han, *An implicit algorithm of solving nonlinear filtering problems*, Comm. Comput. Phys., 16:382–402, 2014. 1

[3] F. Bao, Y. Cao, X. Han, and J. Li, *Efficient particle filtering for stochastic Korteweg-de Vries equations*, Stoch. Dyn., 17(2):1750008, 2017. 1

[4] F. Bao, Y. Cao, A. Meir, and W. Zhao, *A first order fully discretized numerical algorithm for backward doubly stochastic differential equations*, SIAM/ASA J. Uncertain. Quantif., 4(1):413–445, 2016. 1

[5] F. Bao, Y. Cao, C. Webster, and G. Zhang, *A hybrid sparse grid solution of Zakai equation for nonlinear filtering problems*, SIAM/ASA J. Uncertain. Quantif., 2(1):784–804, 2014. 1

[6] F. Bao, Y. Cao, and W. Zhao, *Numerical solutions for forward backward doubly stochastic differential equations and Zakai equations*, Int. J. Uncertain. Quantif., 1(4):351–367, 2011. 2.2, 3.2, 3.2

[7] F. Bao, Y. Cao, and W. Zhao, *A first order semi-discrete algorithm for backward doubly stochastic differential equations*, Discrete Contin. Dyn. Syst. Ser. B, 5(2):1297–1313, 2015. 1

[8] F. Bao, Y. Cao, and W. Zhao, *A backward doubly stochastic differential equation approach for nonlinear filtering problems*, Comm. Comput. Phys., 23(5):1573–1601, 2018. 1

[9] F. Bao and V. Maroulas, *Adaptive meshfree backward SDE filter*, SIAM J. Sci. Comput., 39(6):A2664–A2683, 2017. 1

[10] R.C. Bucy, *Nonlinear filtering theory*, IEEE Trans. Automat. Control, 10:198–199, 1965. 1

[11] A.J. Chorin, M. Morzfeld, and X. Tu, *Implicit particle filters for data assimilation*, Commun. Appl. Math. Comput. Sci., 5(2):221–240, 2010. 1

[12] D. Crisan, *Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation*, Ann. Probab., 31(2):693–718, 2003. 1

[13] D. Crisan and B. Rozovski, *The Oxford Handbook of Nonlinear Filtering*, Oxford University Press, 2011. 1

[14] A. Doucet and A.M. Johansen, *A tutorial on particle filtering and smoothing: Fifteen years later*, in D. Crisan and B. Rozovski (eds.), The Oxford Handbook of Nonlinear Filtering, Oxford University Press, 2011. 1

[15] J. Duník, M. Šimandl, and O. Straka, *Unscented Kalman filter: aspects and adaptive setting of scaling parameter*, IEEE Trans. Automat. Control, 57(9):2411–2416, 2012. 1

[16] I.V. Girsanov, *On transforming a class of stochastic processes by absolutely continuous substitution of measures*, Teor. Verojatnost. i Primenen., 5:314–330, 1960. 2.1

[17] E. Gobet, G. Pagès, H. Pham, and J. Printems, *Discretization and simulation of the Zakai equation*, SIAM J. Numer. Anal., 44(6):2505–2538 (electronic), 2006. 1

[18] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*, IEEE Proc.-F, 140(2):107–113, 1993. 1

[19] Y. Hu, G. Kallianpur, and J. Xiong, *An approximation for the Zakai equation*, Appl. Math. Optim., 45(1):23–44, 2002. 1

[20] S.J. Julier and J.J. LaViola, *On Kalman filtering with nonlinear equality constraints*, IEEE Trans. Signal Process., 55(6):2774–2784, 2007. 1

[21] G. Kallianpur and C. Striebel, *Estimation of stochastic systems: Arbitrary system process with additive white noise observation errors*, Ann. Math. Statist., 39:785–801, 1968. 2.1

[22] G. Kallianpur and C. Striebel, *Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes*, Teor. Verojatnost. i Primenen, 14:597–622, 1969. 1, 2.1

[23] R.E. Kalman and R.S. Bucy, *New results in linear filtering and prediction theory*, Trans. ASME Ser. D.J. Basic Engrg., 83:95–108, 1961. 1

[24] K. Kang, V. Maroulas, I. Schizas, and F. Bao, *Improved distributed particle filters for tracking in a wireless sensor network*, Comput. Stat. Data Anal., 117:90–108, 2018. 1

[25] H.J. Kushner, *Dynamical equations for optimal nonlinear filtering*, J. Diff. Eqs., 2:179–190, 1967. 1

[26] G.N. Milstein and M.V. Tretyakov, *Monte Carlo methods for backward equations in nonlinear filtering*, Adv. Appl. Probab., 41(1):63–100, 2009. 1

[27] M. Morzfeld, X. Tu, E. Atkins, and A.J. Chorin, *A random map implementation of implicit filters*, J. Comput. Phys., 231(4):2049–2066, 2012. 1

[28] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Sixth Edition, Universitext. Springer-Verlag, Berlin, 2003. 2.1

[29] E. Pardoux, *Stochastic particle differential equations and filtering of diffusion processes*, Stochastics, 3:127–167, 1979. 1

[30] É. Pardoux and P. Protter, *A two-sided stochastic integral and its calculus*, Probab. Theory Relat. Fields, 76(1):15–49, 1987. 2.2

[31] É. Pardoux and S. Ge Peng, *Backward doubly stochastic differential equations and systems of quasilinear SPDEs*, Probab. Theory Relat. Fields, 98(2):209–227, 1994. 1, 2.2, 2.2, 2.2, 3.1, 3.2, 3.2

[32] A.N. Shiryaev, *On stochastic equations in the theory of conditional Markov processes*, Theor. Probab. Appl., 11:179–184, 1966. 1

[33] R.I. Stratonovich, *Conditional Markov processes*, Theor. Probab. Appl., 5:156–178, 1960. 1

[34] M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 11:230–243, 1969. 1, 2.1, 3.2